

## RESEARCH ARTICLE

10.1002/2014JD022411

## Key Points:

- Satellite informed fluxes are associated with large uncertainties
- Flux differences are difficult to discern from natural atmospheric variability
- Models continue to underestimate NH sinks

## Correspondence to:

L. E. Ott,  
lesley.e.ott@nasa.gov

## Citation:

Ott, L. E., et al. (2015), Assessing the magnitude of CO<sub>2</sub> flux uncertainty in atmospheric CO<sub>2</sub> records using products from NASA's Carbon Monitoring Flux Pilot Project, *J. Geophys. Res. Atmos.*, 120, doi:10.1002/2014JD022411.

Received 7 AUG 2014

Accepted 30 NOV 2014

Accepted article online 2 DEC 2014

# Assessing the magnitude of CO<sub>2</sub> flux uncertainty in atmospheric CO<sub>2</sub> records using products from NASA's Carbon Monitoring Flux Pilot Project

Lesley E. Ott<sup>1</sup>, Steven Pawson<sup>1</sup>, George J. Collatz<sup>2</sup>, Watson W. Gregg<sup>1</sup>, Dimitris Menemenlis<sup>3</sup>, Holger Brix<sup>4,5</sup>, Cecile S. Rousseaux<sup>1,6</sup>, Kevin W. Bowman<sup>3</sup>, Junjie Liu<sup>3</sup>, Annmarie Eldering<sup>3</sup>, Michael R. Gunson<sup>3</sup>, and Stephan R. Kawa<sup>7</sup>
<sup>1</sup>Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, <sup>2</sup>Biospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, <sup>3</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA, <sup>4</sup>University of California, Los Angeles, California, USA, <sup>5</sup>Now at Helmholtz-Zentrum Geesthacht, Institute of Coastal Research, Geesthacht, Germany, <sup>6</sup>Universities Space Research Association, Greenbelt, Maryland, USA, <sup>7</sup>Atmospheric Chemistry and Dynamics Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

**Abstract** NASA's Carbon Monitoring System Flux Pilot Project (FPP) was designed to better understand contemporary carbon fluxes by bringing together state-of-the-art models with remote sensing data sets. Here we report on simulations using NASA's Goddard Earth Observing System Model, version 5 (GEOS-5) which was used to evaluate the consistency of two different sets of observationally informed land and ocean fluxes with atmospheric CO<sub>2</sub> records. Despite the observation inputs, the average difference in annual terrestrial biosphere flux between the two land (NASA Ames Carnegie-Ames-Stanford-Approach (CASA) and CASA-Global Fire Emissions Database version 3 (GFED)) models is 1.7 Pg C for 2009–2010. Ocean models (NASA's Ocean Biogeochemical Model (NOBM) and Estimating the Circulation and Climate of the Ocean Phase II (ECCO2)-Darwin) differ by 35% in their global estimates of carbon flux with particularly strong disagreement in high latitudes. Based upon combinations of terrestrial and ocean fluxes, GEOS-5 reasonably simulated the seasonal cycle observed at Northern Hemisphere surface sites and by the Greenhouse gases Observing SATellite (GOSAT) while the model struggled to simulate the seasonal cycle at Southern Hemisphere surface locations. Though GEOS-5 was able to reasonably reproduce the patterns of XCO<sub>2</sub> observed by GOSAT, it struggled to reproduce these aspects of Atmospheric Infrared Sounder observations. Despite large differences between land and ocean flux estimates, resulting differences in atmospheric mixing ratio were small, typically less than 5 ppm at the surface and 3 ppm in the XCO<sub>2</sub> column. A statistical analysis based on the variability of observations shows that flux differences of these magnitudes are difficult to distinguish from inherent measurement variability, regardless of the measurement platform.

## 1. Introduction

Major weaknesses still exist in our understanding of the processes that control atmospheric carbon dioxide (CO<sub>2</sub>) concentrations and as a result, our ability to simulate and predict changes in the Earth system. While the magnitude of the global growth rate of atmospheric CO<sub>2</sub> is well constrained by surface observations [e.g., Conway *et al.*, 1994], attributing its changes to specific processes remains a challenge. The Atmospheric Carbon Cycle Inversion Intercomparison (TransCom 3) compared estimates of the carbon budget produced by a variety of inverse models. Their results indicated the presence of a larger Northern Hemisphere (NH) land carbon sink than represented in the a priori flux distribution assumed for the experiment, though the results were strongly influenced by differences in model transport fields [Gurney *et al.*, 2002]. The precise location, cause of, and variability of this missing carbon sink remain poorly understood despite the insights provided by intercomparison studies. In addition to uncertainty in the processes governing the contemporary carbon budget, there is evidence that natural land and ocean carbon sinks have decreased over the course of the twentieth century resulting in an increase in the

fraction of anthropogenic emissions remaining in the atmosphere. However, assessment of changes in the airborne fraction remains controversial [Le Quéré *et al.*, 2009; Knorr, 2009].

Model estimates of land and ocean carbon flux are important because they are one of the main ways to understand the underlying processes governing carbon storage and exchange. These models facilitate an enhanced understanding that may one day contribute to improved abilities to predict changes in the global carbon budget. A number of attempts have been made to compare estimates of carbon flux from different land models. *Randerson et al.* [2009] compared two biogeochemistry models using a common modeling framework and demonstrated that global carbon sinks differed by a factor of 2 during the 1990s. *Schaefer et al.* [2012] and *Raczka et al.* [2013] compared estimates of gross primary production (GPP) and carbon balance among terrestrial ecosystem models over North America and found large differences. *Schwalm et al.* [2010] compared output from 22 terrestrial biosphere models with data from flux towers in North America and found that models' ability to reproduce observed monthly net ecosystem exchange was poor, though performance was better at forested sites than nonforested sites. Most recently, *Huntzinger et al.* [2012] compared 19 terrestrial biosphere models over North America and found large differences in estimates of GPP (between 12.2 and 32.9 Pg C yr<sup>-1</sup>) and heterotrophic respiration (5.6 to 13.2 Pg C yr<sup>-1</sup>) among models with smaller differences in net ecosystem production (−0.7 to 2.2 Pg C yr<sup>-1</sup>). These results underscore the continued uncertainty in land flux processes. Observational estimates of net ecosystem exchange collected worldwide at tower stations provide valuable information to constrain carbon fluxes at local scales [e.g., *Lafleur et al.*, 2003; *Hollinger et al.*, 2004; *Winderlich et al.*, 2014]. However, these observations are difficult to translate to realistic global estimates because of the spatial heterogeneity of vegetation and limited flux sampling locations.

While ocean carbon fluxes are considered to be reasonably well constrained by the observational database compiled by *Takahashi et al.* [2002, 2009], certain regions are poorly sampled and models are still needed to understand and predict how air-sea exchange processes might evolve in the future. As is the case among land models, wide disparity exists among model-derived ocean flux estimates. *Doney et al.* [2004] showed that errors in model physical processes complicated efforts to compare ocean biogeochemical fields among models. *Najjar et al.* [2007] documented large circulation differences among 12 global ocean models that influenced export of dissolved organic matter, particularly to the Southern Ocean. *Popova et al.* [2012] compared five ocean models in the arctic and found that, while primary production was generally consistent among models, the models disagreed over the relative importance of the processes governing production. Uncertainties in the ocean carbon uptake are especially important because they are directly propagated into the inference of the global terrestrial sink as used by the Global Carbon Project [Le Quéré *et al.*, 2013].

It is important to design an Earth observing system that will enable reduction of uncertainties in carbon fluxes and enhance our ability to predict changes in the carbon climate system. A number of components of the carbon cycle are currently constrained by different types of observations. The Moderate Resolution Imaging Spectroradiometer (MODIS) instruments aboard NASA's Terra and Aqua satellites provide information about vegetation characteristics and fire. Information from the advanced very high resolution radiometer (AVHRR) instruments aboard NOAA satellites provide a longer term record of vegetation evolution. The Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Suomi National Polar-orbiting Partnership (Suomi NPP), and MODIS all provide information on ocean color and productivity, while MODIS and geostationary weather satellites observe the physical ocean state. Weather satellites enhance our understanding of the carbon cycle by providing information on atmospheric circulation that is necessary for interpreting atmospheric CO<sub>2</sub> observations. In addition, these weather satellites inform hydrology and temperature estimates used by terrestrial biosphere models to estimate flux. Though these remote sensing products inform a number of model-based land and ocean carbon flux estimates, substantial uncertainty remains as documented in the numerous studies cited above.

Atmospheric CO<sub>2</sub> observations provide an important constraint on carbon fluxes but are limited in space and time. NOAA maintains a network of surface sampling sites in remote locations worldwide where trace gas measurements are conducted several times per week as well as a smaller number of observatories and tower sites that provide continuous CO<sub>2</sub> data. Ground-based column CO<sub>2</sub> observations are currently collected at a small number of stations as part of the Total Carbon Column Observing Network (TCCON). *Chevallier et al.* [2011] demonstrated the utility of column CO<sub>2</sub> data for inferring regional carbon budgets for the first time

using data from this sparse network. Several satellite data sets are available from the Atmospheric Infrared Sounder (AIRS) instrument aboard Aqua and the Japanese Greenhouse gases Observing SATellite (GOSAT), and, prior to May 2010, the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) instrument aboard Environmental Satellite. The impact of satellite CO<sub>2</sub> data is limited by the inability of these instruments to observe in the presence of clouds and aerosols (GOSAT) and low sensitivity to near surface CO<sub>2</sub> mixing ratios (AIRS). Additionally, the observing strategies of GOSAT and SCIAMACHY only permit observations during daylight hours. The use of satellite CO<sub>2</sub> observations for flux inference is still an evolving field of research. *Chevallier et al.* [2009] demonstrated that AIRS data could improve flux estimates, but that these data alone performed worse in inversions when compared with surface flask data. Several recent studies have shown that GOSAT data have the potential to reduce uncertainty in a priori flux estimates [e.g., *Basu et al.*, 2014; *Chatterjee et al.*, 2013; *Guerlet et al.*, 2013; *Maksyutov et al.*, 2013], but the results can be strongly impacted by systematic errors contained in the data as noted by *Basu et al.* [2013] and *Chevallier et al.* [2014].

NASA's Carbon Monitoring System (CMS) Flux Pilot Project (FPP) was designed to better understand the observational constraints placed on carbon fluxes by bringing together state-of-the-art models with a wealth of remote sensing resources. "Bottom up" surface flux estimates were computed by two land and two ocean models for the 2009–2010 period using a consistent set of meteorology input from NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA) [*Rienecker et al.*, 2011]. These results were propagated forward in the atmosphere using NASA's Goddard Earth Observing System Model, version 5 (GEOS-5) constrained by MERRA analyzed meteorological fields. Unlike other model intercomparison projects that span a much larger range of model variables, our intention is not to assess the full range of flux uncertainty and variability. This work brings together a small group of model estimates of flux and atmospheric transport that are informed by remote sensing data sets and draws upon the expertise of the creators of those data sets to assess flux differences and their manifestation in atmospheric CO<sub>2</sub> records. Here we report on the results of these simulations with a focus on understanding how uncertainty in observationally informed land and ocean flux estimates propagates into atmospheric CO<sub>2</sub> and characterizing how differences in flux processes might be observed in the atmosphere by existing platforms. This work complements the application of "top-down" approaches used to infer the surface fluxes given satellite observations [*Liu et al.*, 2014] by analyzing model performance and sensitivity to the assumption of flux in greater detail. Section 2 of this paper describes the models used in this work, while section 3 of this paper describes the comparison with different CO<sub>2</sub> data sets. Section 4 provides a summary and conclusions drawn from this work.

## 2. Model and Data Background

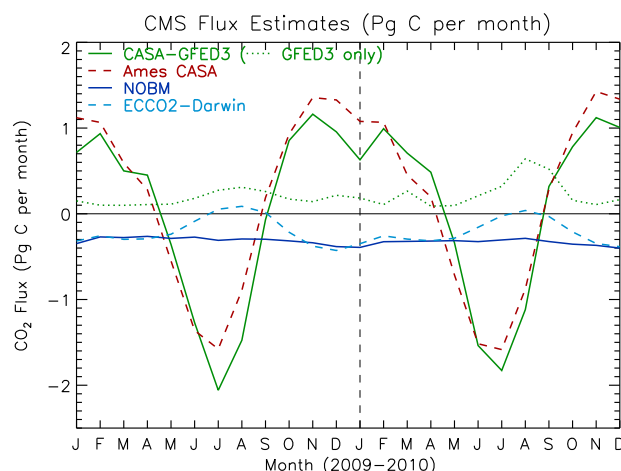
The CMS FPP modeling framework includes land fluxes from the Carnegie-Ames-Stanford-Approach–Global Fire Emissions Database version 3 (CASA-GFED3) and the NASA Ames CASA model (hereafter referred to as Ames CASA) described in sections 2.1.1 and 2.1.2, respectively. Ocean model fluxes come from NASA's Ocean Biogeochemical Model (NOBM; section 2.2.1) and the ECCO2-Darwin model (section 2.2.2). These fluxes were transported using the GEOS-5 Atmospheric General Circulation Model (AGCM) constrained by the MERRA (section 2.3). Model output was sampled at the times and locations of surface, TCCON, and satellite observations as described in section 3 in order to assess overall model performance and the ability of different observational data sets to discern between land and ocean flux estimates.

### 2.1. Land Biosphere Flux Estimates

#### 2.1.1. CASA-GFED3

CASA-GFED3 derives from *Potter et al.* [1993], diverging in development since *Randerson et al.* [1996]. CASA is a light use efficiency type model: net primary production (NPP) is expressed as the product of photosynthetically active solar radiation, a light use efficiency parameter, scalars that capture temperature and moisture limitations, and fractional absorption of solar radiation by the vegetation canopy (FPAR). This latter variable is derived from satellite data.

Fire parameterization was incorporated into the model by *van der Werf et al.* [2004], producing CASA-GFED, and the model has undergone several revisions [*van der Werf et al.*, 2006, 2010] leading to its most recent version CASA-GFED3. Input data sets include air temperature, precipitation, incident solar radiation, a soil



**Figure 1.** Global land and ocean flux estimates produced as part of NASA's CMS FPP for 2009 and 2010 in Pg carbon per month. Dark green solid line indicates combined Net Ecosystem Production (NEP) and biomass burning fluxes computed from by the CASA-GFED3 model (dotted dark green line shows biomass burning contribution only). Brown dashed line indicates NEP from the Ames CASA model. Blue lines show ocean flux estimates from NOBM (dark blue solid) and ECCO2-Darwin (light blue dashed).

classification map, and a number of satellite derived products including MODIS vegetation classification, MODIS-based burned area, and AVHRR FPAR.

CASA-GFED3 is run at monthly time steps with  $0.5^\circ$  spatial resolution. As part of the CMS FPP, fluxes were computed using MERRA meteorology and FPAR derived from AVHRR normalized difference vegetation index (NDVI) [Tucker *et al.*, 2005] according to the procedure of Los *et al.* [2000]. The original 8 km, biweekly AVHRR NDVI was aggregated up to the monthly,  $0.5^\circ \times 0.5^\circ$  grid by averaging. Model output includes Net Primary Production (NPP), heterotrophic respiration (Rh), and fire emissions from forest, savanna, deforestation, peat, and agriculture. For this project, fire emissions were disaggregated from monthly to daily using the MODIS active

fire products as described in Mu *et al.* [2011]. Monthly NPP and heterotrophic respiration were disaggregated to 3 h time intervals following Olsen and Randerson [2004].

### 2.1.2. NASA Ames CASA

Ames CASA relies on satellite observations of vegetation cover from MODIS as time series inputs to estimate monthly carbon fluxes from terrestrial ecosystems worldwide. All model algorithms, parameter settings, and land cover data sets used in Ames CASA for CMS flux computations have been documented by Potter *et al.* [2007].

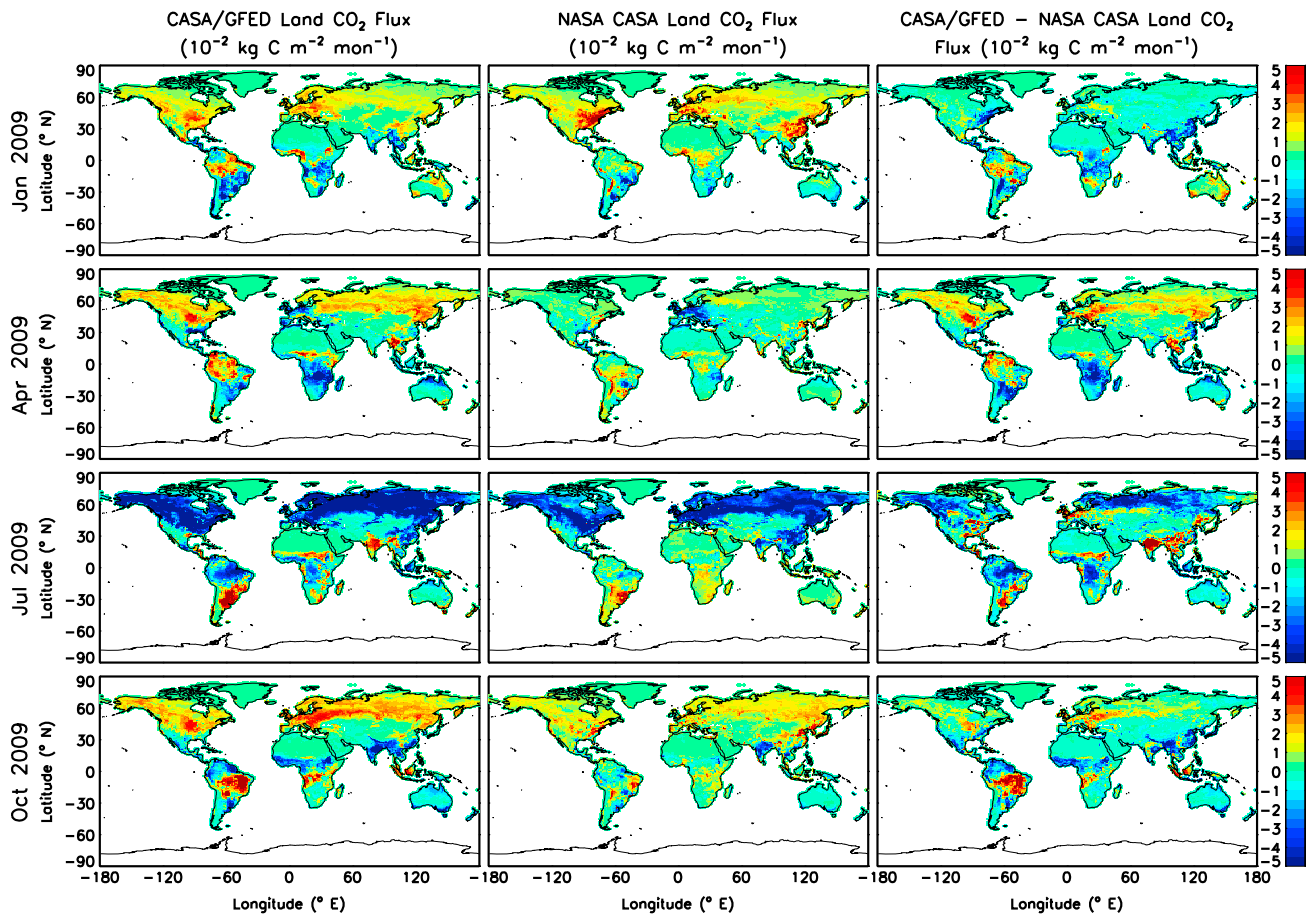
For the CMS FPP computations of net biosphere fluxes of carbon to the atmosphere, the Ames CASA version documented in Potter *et al.* [2007] has been modified to use global  $0.5^\circ$  (latitude/longitude) MODIS enhanced vegetation index (EVI) input data (for the years 2000–2010) generated by aggregating monthly  $0.05^\circ$  ( $\sim 6$  km) values. In addition, in cropland areas, 40% of annual NPP carbon is removed each year from the litter decomposition flux pathways and diverted into harvested food products. This is assumed to be reemitted as a consistent monthly flux (1/12 of the annual cropland harvest carbon total) with a weighted spatial distribution corresponding to the maps of cropland harvest  $\text{CO}_2$  emissions developed by Ciais *et al.* [2007]. Net biosphere fluxes for both CASA-GFED3 and Ames CASA were disaggregated to 3 h time intervals following Olsen and Randerson [2004]. In contrast to the CASA-GFED3 fluxes, biomass burning is not included in the Ames CASA fluxes computed as part of the CMS FPP. CASA-GFED3 does not include the crop redistribution incorporated in Ames CASA flux estimates.

The Ames CASA fluxes produced as part of the CMS FPP differ from a more recent version of Ames CASA described in Potter *et al.* [2012] both in the use of different meteorological driver data and in the methodology of process representation.

### 2.1.3. Differences Between Land Flux Estimates

Figure 1 shows time series of monthly global flux totals computed by the two land models. Ames CASA estimates a weaker land sink (driven by NH summer) and a stronger source during NH winter compared to CASA-GFED3. These differences result in substantially different annual flux estimates for both years; during 2009 (2010), Ames CASA estimates that the terrestrial biosphere is a 2.5 (2.1) Pg C source, while CASA-GFED3 calculates a smaller 0.34 (1.2) Pg C source. The average difference in annual land flux between the two models is 1.7 Pg C. For reference, the uncertainty in land flux between these two models is nearly 20% of the annual mean fossil fuel emissions for the 2009 to 2010 period (9 Pg C) [Boden *et al.*, 2011]. GFED3 fluxes for 2009 (2010) indicate that fires are a global 2.1 (2.9) Pg C source of carbon to the atmosphere, while Ames CASA fluxes considered here do not include the effects of fire. In both models, the land biosphere is a source of





**Figure 2.** Monthly land flux estimates from the (left) CASA-GFED3 and (middle) Ames CASA models and their difference (right, CASA-GFED3 minus Ames CASA) for January, April, July, and October of 2009. Units are  $10^{-2}$  kg carbon per  $m^2$  per month.

carbon to the atmosphere rather than a sink as indicated by multimodel inversion studies [e.g., Gurney *et al.*, 2002]. Using observed atmospheric  $CO_2$  increases, fossil fuel emission inventories, and ocean models, Le Quére *et al.* [2013] estimate the magnitude of the land carbon sink as 3.3 (2.55) Pg C during 2009 (2010). This may be due to the fact that neither model includes several processes thought to be important in explaining the global land sink including  $CO_2$  fertilization [e.g., Bellassen *et al.*, 2011; Vanuytrecht *et al.*, 2011; Piao *et al.*, 2013; Los, 2013], nitrogen deposition [e.g., Esser *et al.*, 2011; Bala *et al.*, 2013; Fleischer *et al.*, 2013; Gerber *et al.*, 2013], and land use history [e.g., Sentman *et al.*, 2011]. In addition to differences in the magnitude of the land biosphere flux estimates, the Ames CASA and CASA-GFED3 models indicate differences in the phasing of the global seasonal cycle in terrestrial carbon flux to the atmosphere. Ames CASA diagnoses an earlier pattern of drawdown during NH spring and an earlier transition from sink to source during NH autumn in 2009.

Figure 2 shows the geographic distribution of land fluxes computed by CASA-GFED3 and Ames CASA for 4 months during 2009. During January, CASA-GFED3 indicates greater drawdown of  $CO_2$  over the Southern Hemisphere (SH) regions of South America and Africa than does Ames CASA. Fire emissions in equatorial Africa and Australia are also evident in CASA-GFED3 while absent in Ames CASA. In the NH, Ames CASA estimates greater release of  $CO_2$  to the atmosphere along the east coasts of the United States and China. In April, Ames CASA fluxes are generally weak in most locations as the model transitions between winter and summer seasons; areas of western Europe have transitioned from a source in January to a sink of  $CO_2$  in April. In contrast, CASA-GFED3 patterns are generally similar in April and January but fluxes from the land to the atmosphere are slightly larger at most NH locations. During July, CASA-GFED3 estimates greater uptake of  $CO_2$  at most NH locations, particularly at high latitudes, than does Ames CASA. This contrast is also evident over equatorial South America while in midlatitude regions of the continent, CASA-GFED3 indicates greater release

of CO<sub>2</sub> to the atmosphere. The models also differ strongly in their assessment of fluxes over India and Southeast Asia during July with CASA-GFED3 indicating a substantial source and Ames CASA indicating a sink. During October, Ames CASA fluxes are generally much weaker than CASA-GFED3 fluxes. Strong fire activity over South America is evident in CASA-GFED3 during October and not included in Ames CASA flux estimates. Differences between the spatial distributions of Ames CASA and CASA-GFED3 fluxes for 2010 (not shown) are qualitatively similar though 2010 was characterized by larger fire emissions in CASA-GFED3.

Both Figures 1 and 2 comparing CASA-GFED3 and Ames CASA show greater uptake in the Northern Hemisphere summer and less emissions during winter from CASA-GFED3. Sensitivity experiments using different input data sets showed that this difference is largely due to the source of FPAR information in the two models. Ames CASA uses EVI as a surrogate for FPAR [Potter *et al.*, 2012], while CASA-GFED3 derives FPAR from the AVHRR-based Global Inventory Modeling and Mapping Studies NDVI [Pinzon and Tucker, 2014] according to Los *et al.* [2000]. FPAR derived from NDVI is generally higher during the growing season in temperate and boreal Northern Hemisphere contributing to the larger uptake during the northern growing season in CASA-GFED3 relative to the Ames CASA model, which dominates the global seasonal cycle shown in Figure 1. The earlier phasing of the seasonal cycle and the larger emissions in the winter shown for Ames CASA arises because Ames CASA uses a soil temperature parameterization [Potter, 1997] that produces delayed respiration in the spring while increasing respiration in the autumn and early winter compared to CASA-GFED3 which uses air temperature to control relative humidity.

## 2.2. Ocean Flux Estimates

### 2.2.1. NOBM

Global ocean carbon dynamics are simulated by the NOBM. It is a three-dimensional representation of coupled circulation, biogeochemical, and radiative processes in the global oceans [Gregg *et al.*, 2003; Gregg and Casey, 2007]. The biogeochemical processes model contains four phytoplankton groups, four nutrient groups, a single herbivore group, and three detrital pools. The phytoplankton groups differ in maximum growth rates, sinking rates, nutrient requirements, and optical properties. Three detrital pools provide for storage of organic material, sinking, and eventual remineralization back to usable nutrients. Radiative transfer calculations provide the underwater irradiance fields necessary to drive growth of the phytoplankton groups and interact with the heat budget. Carbon cycling involves dissolved organic carbon (DOC) and dissolved inorganic carbon [Gregg *et al.*, 2013]. DOC has sources from phytoplankton, herbivores, and carbon detritus, and a sink to dissolved inorganic carbon (DIC). DIC has sources from phytoplankton, herbivores, carbon detritus, and DOC, and communicates with the atmosphere, which can be either a source or sink. The ecosystem sink for DIC is phytoplankton, through photosynthesis. This represents the biological pump portion of the carbon dynamics. The solubility pump portion is represented by the interactions among temperature, alkalinity (parameterized as a function of salinity), silica, and phosphate (parameterized as a function of nitrate). The alkalinity/salinity parameterization utilizes the spatial variability of salinity in the model adjusted to mean alkalinity from the Ocean Model Intercomparison Project (OCMIP) [Orr *et al.*, 2001]. The calculations for the solubility pump follow the standards set by the OCMIP.

NOBM's domain spans from  $-84^{\circ}$  to  $72^{\circ}$  latitude in increments of  $1.25^{\circ}$  longitude by  $2/3^{\circ}$  latitude, including only open ocean areas where bottom depth is greater than 200 m. The model contains 14 vertical layers, in quasi-isopycnal coordinates and is driven by MERRA monthly mean wind stress, sea surface temperature, and shortwave radiation fields. MODIS-Aqua chlorophyll data were assimilated to represent the sum of all phytoplankton components in the model.

As part of the CMS FPP, daily NOBM ocean pCO<sub>2</sub> fields were used as input into GEOS-5 and ocean atmosphere CO<sub>2</sub> fluxes were computed within GEOS-5 using the GEOS-5 atmospheric CO<sub>2</sub> mixing ratios and winds. This online ocean calculation results in CO<sub>2</sub> fluxes calculated at the GEOS-5 physics time step (30 min) that are physically consistent with the evolution of simulated weather systems.

### 2.2.2. ECCO2-Darwin

The ECCO2-Darwin Ocean Carbon Cycle Model is based on a global, eddying, ocean and sea ice configuration of the Massachusetts Institute of Technology general circulation model [Marshall *et al.*, 1997a, 1997b]. It combines results from two separate projects: the Estimating the Circulation and Climate of the Ocean, Phase II (ECCO2) Project, which provides a data-constrained estimate of the time-evolving physical ocean state, and the Darwin Project, which provides time-evolving ocean ecosystem variables. Together, ECCO2 and Darwin

provide a time-evolving physical and biological environment for carbon biogeochemistry, which is used to compute surface fluxes of carbon at high spatial and temporal resolution.

The ECCO2 model configuration is a cube-sphere grid [Adcroft *et al.*, 2004] with 18 km horizontal grid spacing and 50 vertical levels [Menemenlis *et al.*, 2005a, 2008]. The domain includes all global ocean basins. The ECCO2 model configuration includes a dynamic/thermodynamic sea ice model [Losch *et al.*, 2010; Heimbach *et al.*, 2010]. In a first step, the ECCO2 model configuration was adjusted using a low-dimensional (Green's functions) estimation approach [Menemenlis *et al.*, 2005b]. In a second step, the method of Lagrange multipliers (adjoint method) was used to adjust initial and time-evolving surface boundary conditions [Wunsch and Heimbach, 2007]. Data constraints include sea level anomaly from Jason-1 and the Ocean Surface Topography Mission (OSTM) aboard Jason-2, sea surface temperature from the Advanced Microwave Scanning Radiometer-EOS (AMSR-E), and temperature and salinity profiles from Argo. This adjoint-based ECCO2 solution is used to drive the Darwin ecosystem model.

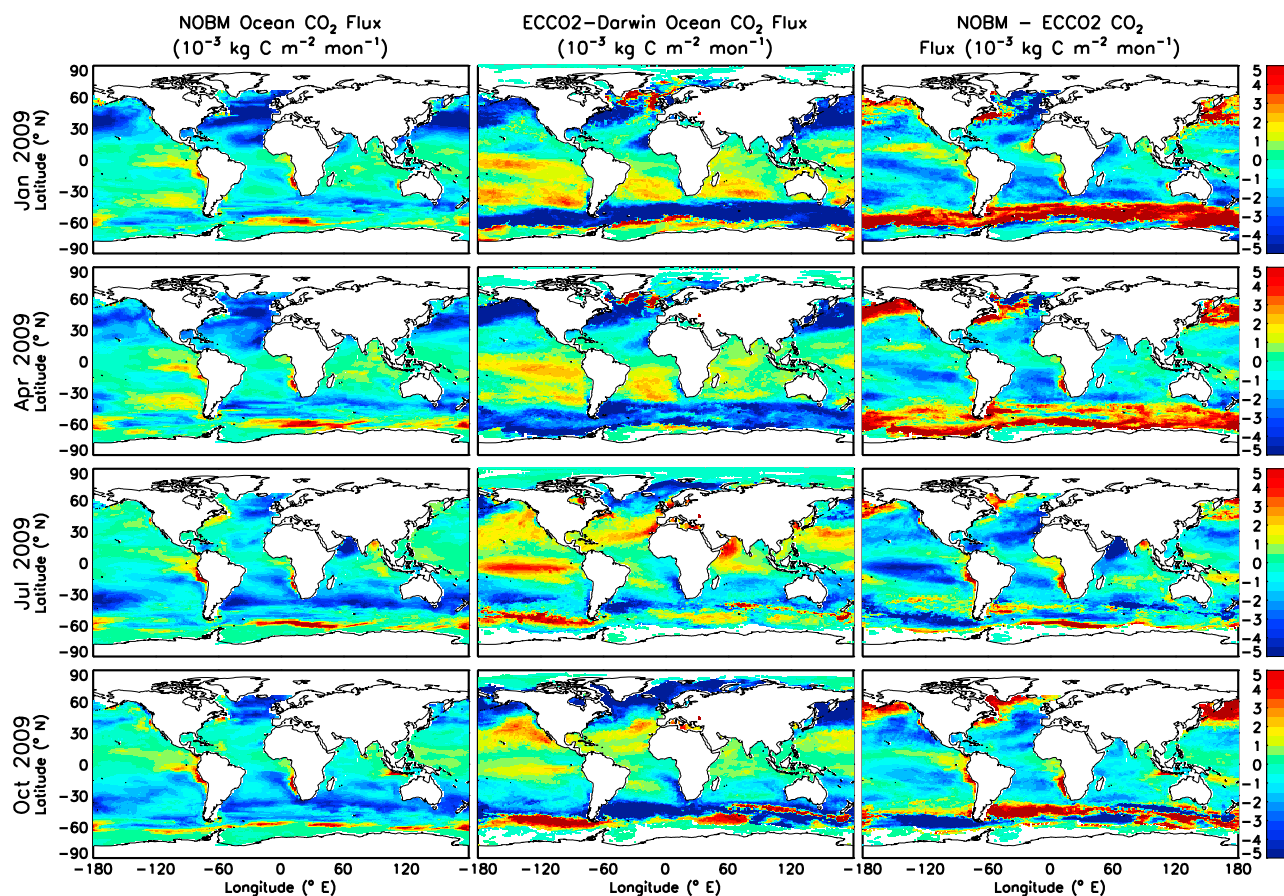
The Darwin Project is an initiative to advance the development and application of novel models of marine microbial communities, identifying the relationships of individuals and communities to their environment, connecting cellular-scale processes to global microbial community structure [Follows *et al.*, 2007; Follows and Dutkiewicz, 2011; Dutkiewicz *et al.*, 2009]. The particular configuration used for the CMS FPP includes five phytoplankton functional types (choices based on results from previous versions of the model) and two zooplankton types. The carbon cycle is explicitly included in this configuration, along with those of nitrogen, phosphorus, iron, silica, oxygen, and alkalinity. The carbonate chemistry follows the simplified model proposed by Follows *et al.* [2006], and air-sea CO<sub>2</sub> exchange is parameterized according to Wanninkhof [1992].

ECCO2-Darwin fluxes produced as part of the CMS FPP are described in more detail in H. Brix *et al.* (Using Green's Functions to initialize and adjust a global, eddying ocean biogeochemistry general circulation model, submitted to *Ocean Modeling*, 2014). GEOS-5 simulations presented here use version 2.1 fluxes. Since the CMS FPP, ECCO2-Darwin fluxes have undergone further development including modification of the piston velocity parameterization.

### 2.2.3. Differences Between Ocean Flux Estimates

On a global basis, NOBM ocean fluxes contain very little seasonal variation while the ECCO2-Darwin fluxes include a seasonal cycle with amplitude of 0.3 Pg C (Figure 1). The small seasonal cycle in NOBM derives from use of seasonally invariant atmospheric pCO<sub>2</sub> data and is only modestly rectified by the online flux calculation used in conjunction with GEOS-5. Both models estimate comparable global fluxes for much of the year but during June, July, August, and September, ECCO2-Darwin fluxes indicate a weaker ocean sink (or slight source) relative to NOBM. NOBM uptake of carbon occurs as a function of enhanced biological activity near the subpolar front, which is driven by the assimilation of satellite chlorophyll observations. As a result, the 2009 (2010) annual ECCO2-Darwin ocean sink is 2.4 (2.6) Pg C compared to 3.7 (4.0) Pg C from NOBM, a difference of 36% (35%). The Global Carbon Budget (GCB) [Le Quéré *et al.*, 2013] estimates a global ocean flux of 2.57 (2.55) Pg C yr<sup>-1</sup> for 2009 (2010). NOBM ocean fluxes indicate a stronger ocean sink than the GCB estimate, while ECCO2-Darwin estimates differ by less than 10% from GCB in both years.

Figure 3 shows geographic distributions of ocean fluxes computed by NOBM and ECCO2-Darwin for 4 months in 2009. During January, differences between model estimates of flux are largest in high-latitude oceans. NOBM estimates a greater Atlantic Ocean sink north of 30°, while ECCO2-Darwin estimates a much stronger ocean sink in the Southern Ocean than is evident in NOBM. The differences are the result of high outgassing in ECCO2-Darwin in the North Atlantic east of Greenland and modest sources to the atmosphere in the Arctic extreme latitudes that are not covered by NOBM. April distributions are generally similar to January, though a decrease in the northern NOBM sink and increase in the ECCO2-Darwin sink relative to January result in spatially inhomogeneous differences north of 30°. During July, ocean flux distributions differ substantially from January; the northern sink in NOBM has weakened as the spring bloom has faded and ocean temperatures have risen and is now evident only in the north Atlantic basin. ECCO2-Darwin's strong Southern Ocean sink has also weakened while the NOBM Southern Ocean sink has become stronger. The spatial patterns are similar with the subpolar frontal zone in location, but NOBM's drawdown is enhanced as described above while ECCO2-Darwin exhibits high outgassing along 60° latitude west of the Drake Passage. During October, the NOBM distribution of fluxes is dominated by a strong sink throughout much of the Southern Hemisphere oceans. ECCO2-Darwin indicates stronger sinks in high-latitude regions



**Figure 3.** Monthly ocean flux estimates from the (left) NOBM and (middle) ECCO2-Darwin models and their difference (right, NOBM minus ECCO2-Darwin) for January, April, July, and October of 2009. Units are  $10^{-3}$  kg carbon per  $m^2$  per month (note that the ocean fluxes shown here are an order of magnitude smaller than the land fluxes shown in Figure 2).

of both hemispheres but with less areal extent relative to NOBM. Ocean models also differ in their estimation of coastal fluxes, particularly off the Atlantic coasts of Africa and North America and the Pacific coast of South America.

In regions below  $30^{\circ}\text{S}$ , ECCO2-Darwin estimates an ocean sink 27% (23%) greater than NOBM during 2009 (2010). In addition to differences in magnitude of the sink in this region which includes the Southern Ocean, the seasonal cycle of ocean fluxes is nearly reversed; NOBM estimates peak drawdown in the sink during SH winter months, while ECCO2-Darwin indicates that the sink is at its weakest during these months. In SH summer, ECCO2-Darwin simulates a maximum in  $\text{CO}_2$  drawdown in the region while NOBM simulates a minimum. The differences are likely related to the model configurations. ECCO2-Darwin employs a dynamic ice model while NOBM utilizes sea ice data. Also, ECCO2-Darwin assimilates local physical conditions in contrast to NOBM, which assimilates satellite chlorophyll. NOBM, whose domain only extends to  $72^{\circ}\text{N}$ , estimates a sink several times weaker than ECCO2 in NH high-latitude oceans. The largest disparities between ocean flux models occur in regions that are rarely observed by ships or are observed only during certain months as shown in Takahashi et al. [2009].

### 2.3. GEOS-5 Earth System Model

The GEOS-5 AGCM has been developed as a flexible tool to represent the atmosphere on a variety of temporal and spatial scales. It is a central component of the GEOS-5 atmospheric data assimilation system [Reinecker et al., 2008], where it is used with  $0.5^{\circ}$  spatial resolution for meteorological analysis and forecasting [Zhu and Gelaro, 2008] including the production of MERRA which spans the period from 1979 to present [Reinecker et al., 2011]. It has also been developed as a tool for studying atmospheric composition and



**Table 1.** Combinations of Surface (Land Biosphere and Oceanic) Fluxes Used in the Model Computations in the Flux Pilot Project<sup>a</sup>

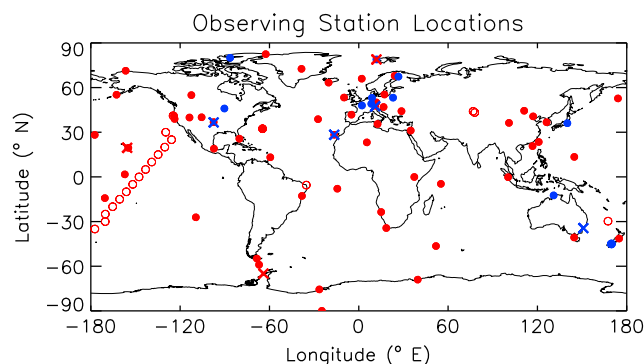
Flux Combination	Land Biosphere	Ocean	Total Land and Ocean Flux (Pg C yr <sup>-1</sup> )	Annual Mean Growth Rate at MBL Sites (ppm yr <sup>-1</sup> )
CG-NO	CASA/GFED-3	NOBM	−3.4 (2009), −2.8 (2010)	2.7
AC-NO	Ames CASA	NOBM	−1.2 (2009), −1.9 (2010)	3.5
CG-ED	CASA/GFED-3	ECCO2-Darwin	−2.1 (2009), −1.4 (2010)	3.3
AC-ED	Ames CASA	ECCO2-Darwin	0.1 (2009), −0.5 (2010)	4.1

<sup>a</sup>Fossil fuel fluxes were identical in all four combinations.

climate. *Ott et al.* [2010] used GEOS-5 to examine the impact of Indonesian biomass burning aerosols on atmospheric circulation, and *Ott et al.* [2011] examined the impact of uncertainty in GEOS-5 convection on global carbon monoxide distributions. Model output has also been used by *Wang et al.* [2009] to compute CO<sub>2</sub> correlations and their impact on inverse modeling.

The AGCM combines the finite volume dynamical core described in *Lin* [2004] with the GEOS-5 column physics package, summarized in *Reinecker et al.* [2008]. The model domain extends from the surface to 0.01 hPa and uses 72 hybrid layers that transition from terrain following near the surface to pure pressure levels above 180 hPa. In this study, the horizontal resolution is 1° × 1.25° (latitude by longitude) and the time step is 30 min for physical computations, with more frequent computations of resolved-scale transport in the dynamical core. Trace gases are transported online in GEOS-5 using the *Lin* [2004] dynamical core for resolved scales, turbulent mixing of CO<sub>2</sub> is performed in the same way as for moisture (using the *Lock et al.* [2000] boundary layer turbulence module), and using the Relaxed-Arakawa Schubert convective scheme [*Moorthi and Suarez*, 1992] to represent convective transport. In the present simulations, transport is constrained with MERRA fields to ensure consistency with observed meteorology.

Land biosphere, biomass burning, fossil fuel, and ocean CO<sub>2</sub> fluxes are prescribed in GEOS-5. For the CMS FPP simulations, GEOS-5 was configured to simulate the emission and transport of several different CO<sub>2</sub> tracers representing differing combinations of land and ocean fluxes described in Table 1. In addition to the land biosphere and ocean fluxes provided as part of the CMS FPP, CO<sub>2</sub> emissions from fossil fuels are taken from the Department of Energy's Carbon Dioxide Information Analysis Center [*Boden et al.*, 2011]. Prior to the target 2009–2010 CMS period, CO<sub>2</sub> tracers were spun up from 2000 to 2008 beginning with a uniform initial condition of 350 ppm to ensure realistic atmospheric distributions. During the spin-up period, land biosphere and biomass burning fluxes from an earlier version of the CASA-GFED model were used in combination with ocean and fossil fuel fluxes from the TransCom Continuous Experiment [*Law et al.*, 2008]. Simulated CO<sub>2</sub> mixing ratios for December 2008 were calculated at the locations of NOAA Earth Science Research Laboratory (ESRL) marine boundary layer stations [*Novelli et al.*, 1992] and compared with observations; on the basis of this comparison a uniform global offset was subtracted from the simulated CO<sub>2</sub> fields to ensure that global average surface CO<sub>2</sub> concentration was representative of atmospheric conditions at the beginning of the CMS period.



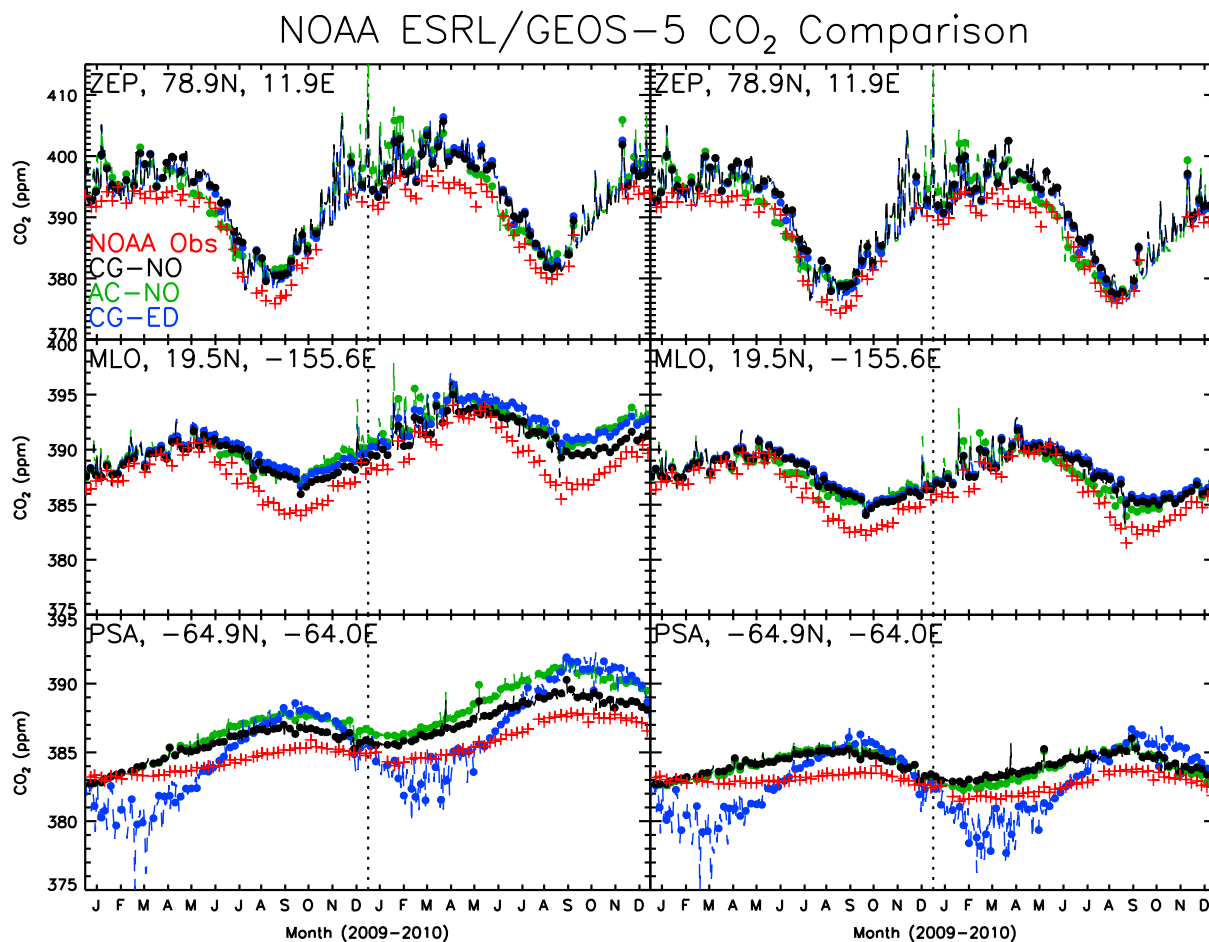
**Figure 4.** Locations of NOAA (red) and TCCON (blue) observing stations in operation during the study period. Open red circles indicate stations that collected an insufficient amount of data in 2009–2010 and thus were not included in Figures 6–8. Red (blue) “cross” marks indicate NOAA (TCCON) stations shown in Figure 5 (Figure 12).

### 3. Comparison With CO<sub>2</sub> Observational Records

#### 3.1. Comparison With Surface CO<sub>2</sub> In Situ Observations

NOAA's Earth System Research Laboratory (ESRL) Carbon Cycle Greenhouse Gases (CCGG) group analyzes samples taken weekly at an international cooperative network of surface observing stations [*Tans et al.*, 1990]. These data have been used to understand both long-term changes and interannual variability of natural carbon sinks [e.g., *Ballantyne et al.*, 2012; *Conway et al.*, 1994]. Sites are typically located in remote locations (Figure 4) so

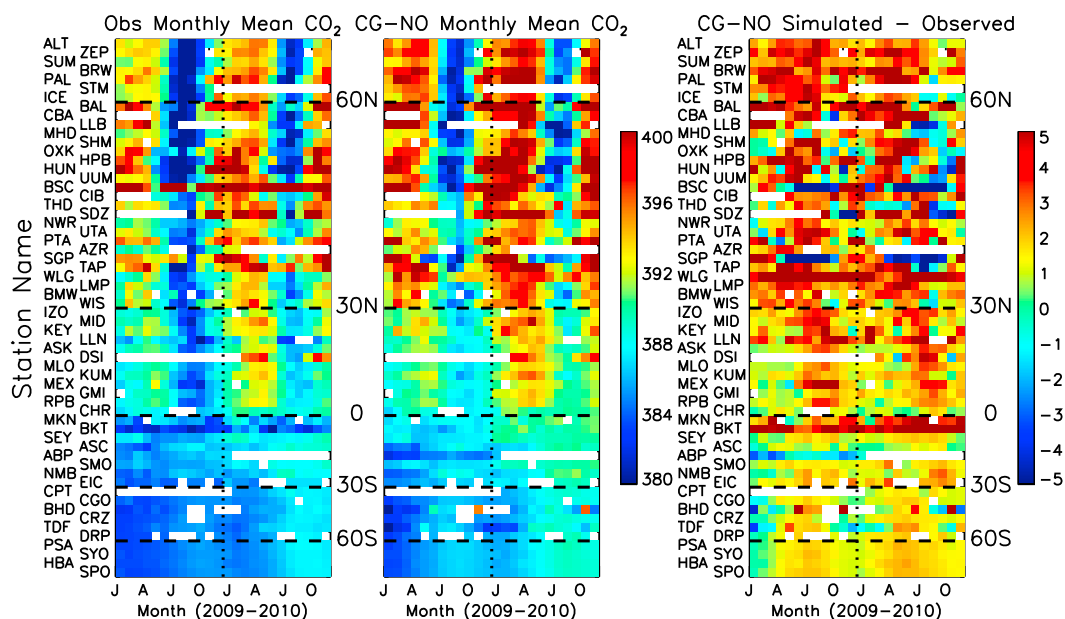




**Figure 5.** (left columns) Simulated and observed  $\text{CO}_2$  at the Ny-Alesund (ZEP; top), Mauna Loa (MLO; middle), and Palmer Station (PSA; bottom). (right columns) Simulated and observed  $\text{CO}_2$  at the same locations when differences in atmospheric growth rate are removed. Red indicates NOAA ESRL observations while black, green, and blue indicates GEOS-5 simulated mixing ratios assuming flux scenarios CG-NO, AC-NO, and CG-ED, respectively (flux scenarios are described in detail in Table 1).

that observations represent the background surface  $\text{CO}_2$  mixing ratios rather than local source and sink influences. More information on specific site locations is provided on the NOAA ESRL web page (<http://www.esrl.noaa.gov/gmd/ccgg/ggrn.php>).

For comparison to surface stations, the GEOS-5 grid cell containing each station location was sampled at the time that an observation was collected and simulated  $\text{CO}_2$  mixing ratios then vertically interpolated to the altitude of the observing station creating model “pseudodata.” Observations and pseudodata are averaged over the course of a day (when multiple observations are present), and daily simulated and observed  $\text{CO}_2$  mixing ratios at Mauna Loa, Ny-Alesund, and Palmer Station (indicated in Figure 4) are shown in Figure 5 for the flux scenarios described in Table 1. One of the primary features evident in this comparison is difference in annual atmospheric growth rate between the flux scenarios. In the beginning of 2009, all scenarios in the GEOS-5 simulations begin from the same  $\text{CO}_2$  mixing ratios, but over time differences in the magnitude of the combined land and ocean carbon flux cause the spread in the ensemble of simulations to grow. While this spread in simulations caused by differing growth rates happens at all stations, it is most evident at Mauna Loa and Palmer Station because of the smaller amplitude of the seasonal cycle at these locations. We calculate growth rates for each flux scenario by first calculating the monthly mean of surface  $\text{CO}_2$  mixing ratio at all CCGG marine boundary layer (MBL) sites during 2009 and 2010. For each month, the 2009 MBL monthly mean is subtracted from the corresponding 2010 value and these differences are averaged over all calendar months to estimate the annual mean growth rate shown in Table 1. Using the same method to calculate the annual mean growth rate from observations yields a value of 2.4 ppm. Flux scenario CG-NO, which combines NOBM and

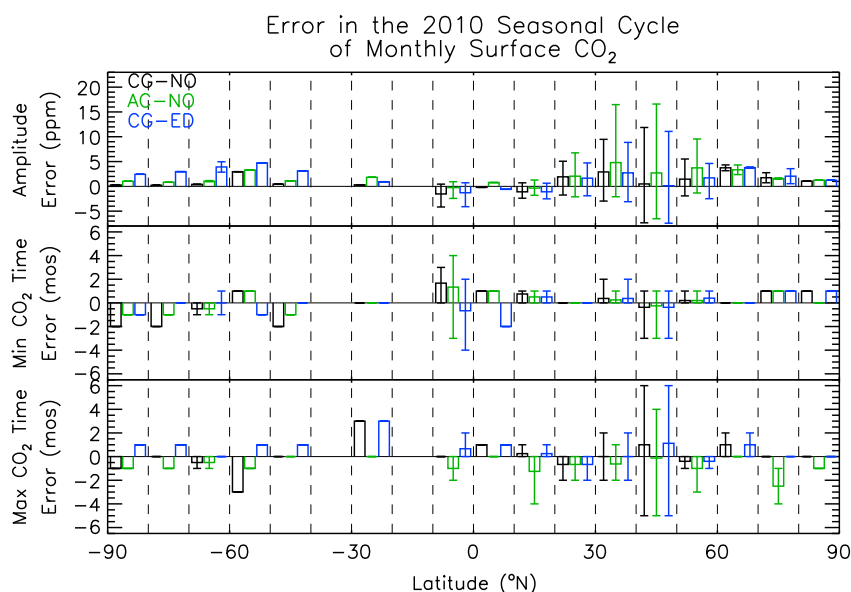


**Figure 6.** Monthly mean CO<sub>2</sub> (ppm) observed at NOAA ESRL stations during (left) 2009–2010 compared with simulated CO<sub>2</sub> assuming flux combination CG-NO ((middle) no detrending applied). (right) The model-observation difference. Stations are oriented from north (top) to south (bottom) on all plots.

CASA-GFED fluxes, provides the most favorable comparison with the observed growth rate. Both the Ames CASA and ECCO-2 fluxes produce weaker natural sinks than flux scenario CG-NO resulting in greater rates of accumulation of CO<sub>2</sub> in the atmosphere for scenarios AC-NO and CG-ED relative to scenario CG-NO.

In order to separate the spatial and temporal differences in flux from the difference in atmospheric growth rate of CO<sub>2</sub>, we temporarily remove the difference in growth rate by subtracting the trends from observations and simulations (Figure 5, right columns only). At Mauna Loa, the detrended time series shows that land biosphere flux differences have a greater impact on simulated mixing ratios than do ocean flux differences. None of the flux scenarios succeeds in reproducing the magnitude of the seasonal cycle observed at Mauna Loa, likely because the NH land sinks in both estimates are too weak or because of deficiencies in model transport. At Ny-Alesund, all model simulations reproduce the seasonal cycle reasonably well though simulations tend to overestimate CO<sub>2</sub>, particularly during spring. Differences in seasonal cycle between the Ames CASA and CASA-GFED fluxes are small but evident at this station in spring and summer months. Ocean flux differences manifest as much smaller mixing ratio differences despite the fact that only one model, ECCO2-Darwin, produces fluxes at this high-northern latitude region. At Palmer Station Antarctica, ocean flux differences are larger than at any other station due to the disparity between model estimates of Southern Ocean flux. NOBM overestimates the magnitude of the observed, weak seasonal cycle by 0.5 ppm, while seasonal variations in the ECCO2-Darwin simulation are much too strong resulting in a 5 ppm overestimate of the seasonal cycle amplitude. The similarity of the CG-NO and AC-NO simulations and large differences between the CG-NO and CG-ED simulations suggests that the seasonal cycle of nearby ocean fluxes more strongly influences the simulation of CO<sub>2</sub> at high-latitude SH stations than does the seasonal cycle of land fluxes.

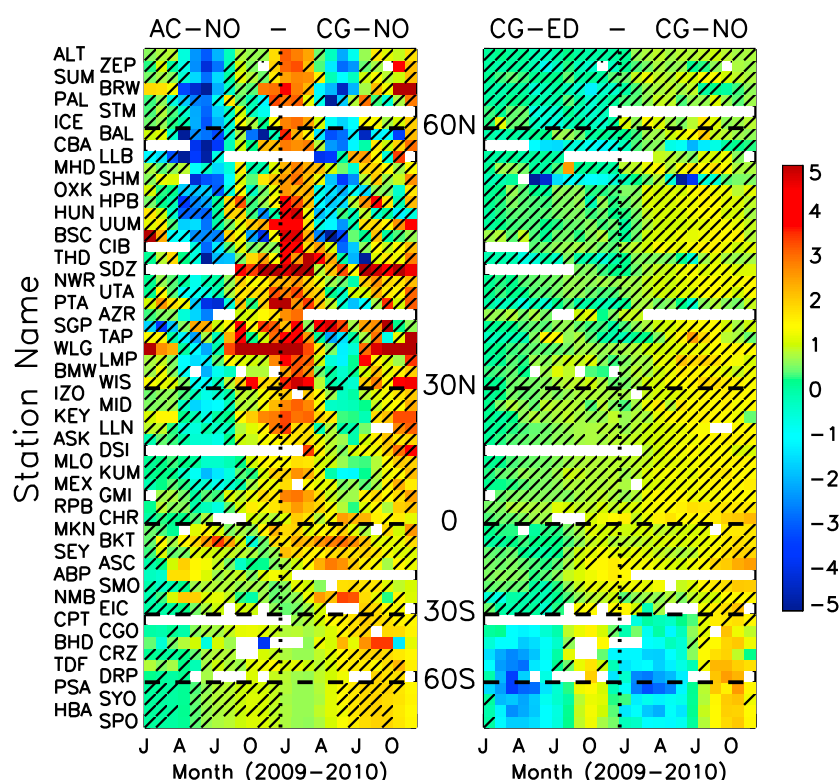
Figure 6 compares observed and simulated (assuming flux combination CG-NO without detrending) monthly mean CO<sub>2</sub> mixing ratios at all NOAA ESRL stations collecting substantial amounts of data during the 2009–2010 period. Observations at NH stations comprise the bulk of the data set and show a strong seasonal cycle due to the influence of the land biosphere. The CG-NO simulation also shows a strong seasonal cycle for most stations north of 30°N. At these locations, the model tends to overestimate CO<sub>2</sub>, particularly during NH spring and summer, because the CASA-GFED3 land sink is too weak. Comparisons with the GCB, presented by *Le Quéré et al.* [2013], indicate an overestimate of global fluxes by the GC-NO flux combination of approximately 3 Pg C which translates into an average, global overestimate in the atmospheric CO<sub>2</sub> of 1.4 ppm assuming that the excess CO<sub>2</sub> were



**Figure 7.** Evaluation of the simulations' ability to represent the seasonal cycle of  $\text{CO}_2$  observed at NOAA surface stations during 2010. (top) The mean error in the magnitude of the seasonal cycle calculated over  $10^\circ$  latitude bins and for flux combinations CG-NO, AC-NO, and CG-ED. The mean error in the month during which the (middle) minimum (bottom, maximum) in the seasonal cycle occurs. Boxes identify mean values while vertical lines indicate maximum and minimum values within each  $10^\circ$  latitude range.

instantly diffused throughout the mass of the atmosphere. At stations north of  $60^\circ\text{N}$ , annual mean  $\text{CO}_2$  is, on average, overestimated by 2.3 ppm during 2010, while stations between the equator and  $60^\circ\text{N}$  are overestimated by an average of 1.3 ppm.  $\text{CO}_2$  at northern high-latitude stations is overestimated by an average of 4 ppm during spring months and 3 ppm during summer, while NH stations south of  $60^\circ\text{N}$  are overestimated by 1.7 ppm during spring and 1.9 ppm during summer. The difference in the magnitude of overestimate between NH high and midlatitude stations may indicate that the CASA-GFED3 land sink underestimate is concentrated in this area or that vertical transport errors are greater at NH high latitudes, resulting in excessive concentration of  $\text{CO}_2$  near the surface. In the SH, model errors average 1.7 ppm during 2010 and show much weaker seasonality.

In Figure 7, we evaluate the ability of GEOS-5 simulations to reproduce the seasonal cycle in  $\text{CO}_2$  observed at ESRL stations during 2010. To calculate seasonal cycle errors, monthly means were calculated from observations and model output after detrending to remove the annual growth rate. We define the amplitude of the seasonal cycle as the difference between the maximum and minimum monthly values. Errors in the months during which minima and maxima in the seasonal cycle occur are also calculated for each station. Means of the amplitude and phasing errors are then calculated over  $10^\circ$  latitude bands and only include stations for which data are available in every calendar month during 2010. At NH midlatitude and high-latitude stations, performance is similar for all model flux combinations. Simulations overestimate the magnitude of the seasonal cycle by a few parts per million on average, generally succeed in estimating the summer month during which the minimum should occur, but have more difficulty in estimating the timing of maximum  $\text{CO}_2$  ratios in the winter. While errors in the amplitude of the seasonal cycle are small at tropical stations, the model given any flux combination tends to struggle reproducing the observed timing of minima and maxima at SH tropical stations. This is likely due to a combination of the weak seasonal cycle in  $\text{CO}_2$  at these locations, errors in the transport of  $\text{CO}_2$  from NH locations, and uncertainty in land fluxes in this region that contains dense vegetation. All simulations considered in this study fail to adequately simulate the seasonal cycle at SH high-latitude stations. Amplitude errors of several parts per million are often larger than the observed seasonal cycle magnitude. In these locations, ECCO2-Darwin fluxes, which indicate a stronger Southern Ocean sink with greater seasonality relative to NOBM, degrade model performance in terms of both seasonal cycle amplitude and the timing of the maximum in surface  $\text{CO}_2$  mixing ratio.



**Figure 8.** Differences in monthly mean  $\text{CO}_2$  mixing ratio (no detrending) at NOAA surface stations due to differing (left) land flux estimates and (right) ocean flux estimates. Diagonal bars indicate instances in which flux differences would not be statistically significant from observation variability.

Despite the substantial flux differences among the land and ocean models shown in Figures 1–3, differences in atmospheric  $\text{CO}_2$  mixing ratio at remote surface stations tend to be quite subtle as evident in the time series comparisons in Figure 5. To quantitatively determine how atmospheric  $\text{CO}_2$  observations might be used to discern between these observationally informed flux estimates, we follow the method of *Huntzinger et al.* [2011] who examined the utility of continuous atmospheric flux measurements at a set of tower locations for constraining biospheric flux variability. In order to test the statistical significance of the difference between simulated pseudoobservations, we apply a chi-square test of variance based on the number of observations per month, the mean squared difference between pairs of pseudoobservations using different flux estimates, and an estimate of expected model-data mismatch [*Huntzinger et al.*, 2011]. Expected model-data mismatch, an estimate of how closely a model could reproduce an observation, is estimated in this study by estimating the variance of observation residuals around a smooth curve fit to the observations at each site, a method used by *Bousquet et al.* [1999] and *Gurney et al.* [2002]. We calculate mean squared difference between simulated  $\text{CO}_2$  pairs by first removing the difference in annual growth rate as shown in Figure 5. Qualitatively, this method allows us to assess how well the small atmospheric signal caused by a difference in assumed flux might be distinguished from a background of strong variability. It is important to note that expected model-data mismatch is calculated solely based on variability in the observations in this study, though several other studies have explored alternate methods of estimating this quantity [e.g., *Michalak et al.*, 2004; *Schuh et al.*, 2010; *Gourdji et al.*, 2010].

Figure 8 shows the difference in monthly mean  $\text{CO}_2$  mixing ratio at NOAA ESRL stations resulting from underlying land and ocean flux differences (no detrending applied). Differences which are not significantly larger than the observation variance at a significance level of 0.05 are indicated by diagonal lines. The largest differences due to land flux are evident at stations in the NH midlatitudes and high latitudes during spring and early summer (2–4 ppm), when Ames CASA indicates an earlier drawdown of  $\text{CO}_2$  by the land biosphere than does CASA-GFED3, and in January through March of 2010 (2–5 ppm), when the larger land to

atmosphere carbon flux in Ames CASA from the latter part of 2009 is most evident. A similar feature is not present in 2009 because it is too close to the initiation of the simulations, but begins to be seen again in December 2010. Land flux differences are typically less than 2 ppm at most remote SH locations. At NH locations, land flux differences manifest primarily as a difference in seasonal cycle, consistent with the global flux differences shown in Figure 1. Localized differences, such as those shown in Figure 2 during April and October, are generally not evident in monthly mean CO<sub>2</sub> mixing ratios at NOAA ESRL stations because of their remote locations. The lack of evidence of fire emissions, present in CASA-GFED3 and not Ames CASA, is striking, particularly at SH locations. Analysis of the model's GFED3 biomass burning tracer shows that at all stations except Ascension Island, deviations from the annual mean are smaller than 1 ppm. At Ascension Island, biomass burning events with greater than 1 ppm influence only occur on a few days.

Even in locations and months where substantial differences in CO<sub>2</sub> mixing ratio due to land flux exist, their magnitude is small enough that they are often difficult to distinguish from variability evident in the observations. In the NH midlatitudes, differences in CO<sub>2</sub> mixing ratio attributable to land flux would not be significantly larger than the observation variance because of the large degree of variability evident in the observational records. Differences at NH high-latitude stations and in the SH are statistically significant but only during certain seasons.

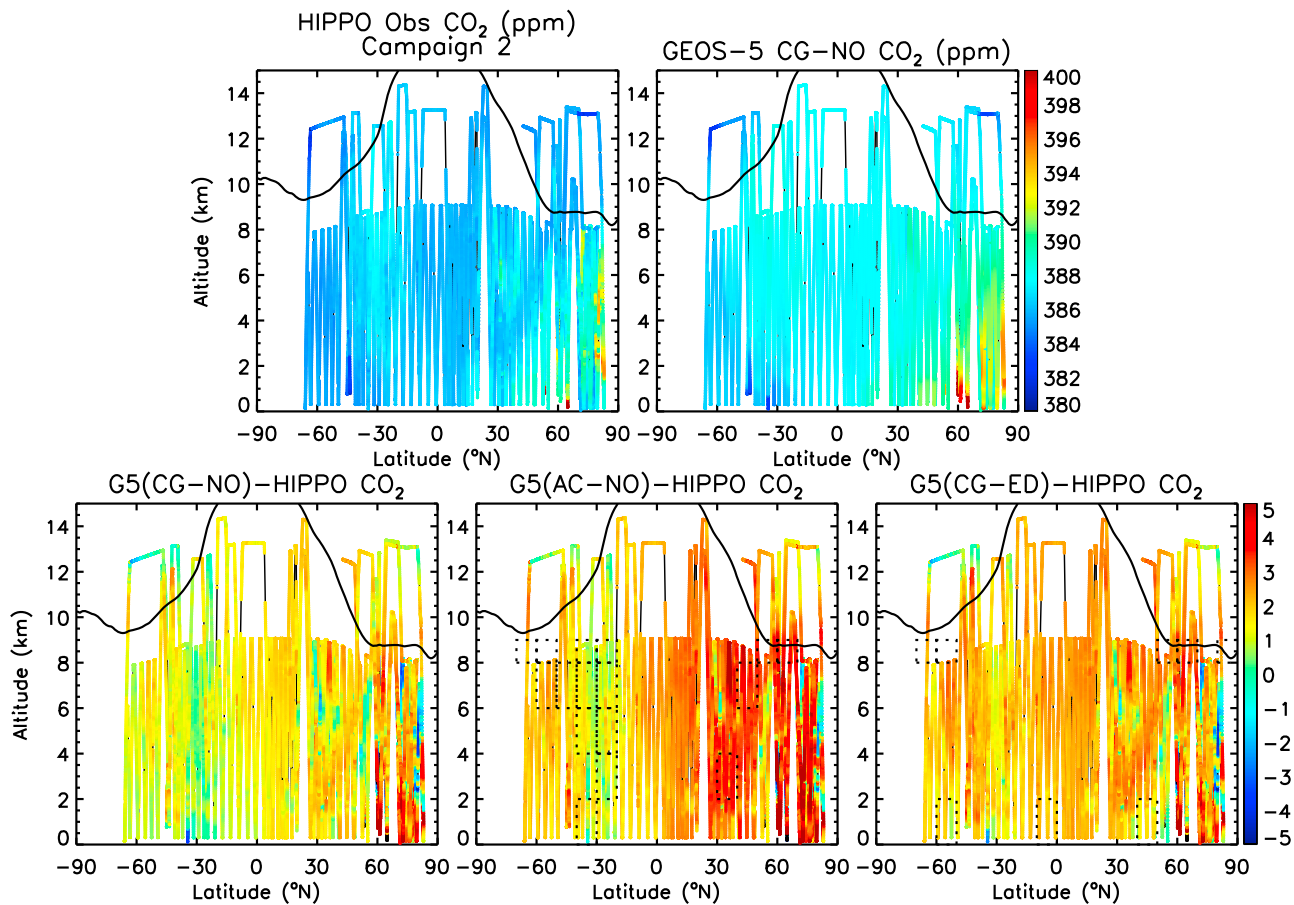
Differences between ocean flux models, which are nearly as large as land flux differences on a global annual basis but exhibit much less seasonality, are even less evident at NOAA ESRL stations. Large differences due to the difference in Southern Ocean sink exist only at stations south of 30°S but are statistically significant in most months. Significant differences between the two ocean models at all stations south of 30°S demonstrate the importance of ocean flux in that region. There, NOBM fluxes are more realistic than ECCO2-Darwin fluxes which result in underestimates in surface CO<sub>2</sub> mixing ratio and errors in the gradient between tropical and SH high-latitude CO<sub>2</sub>. Observations suggest a mean difference between Mauna Loa and SH high-latitude stations during March, April, and May (MAM) of 6.6 ppm. While NOBM fluxes reproduce this spatial gradient fairly well, simulating a mean difference of 6.1 ppm, ECCO2-Darwin fluxes result in a gradient of 8.5 ppm indicating that Southern Ocean drawdown is too large during certain months. Large differences in ocean CO<sub>2</sub> flux are also present in the NH (Figure 3), but the fact that they occur over smaller regions than in the SH results in smaller mixing ratio differences at NH locations. The combination of smaller mixing ratio differences and greater variability due to biosphere flux and fossil fuel emissions in the NH means that NH ocean flux differences are difficult to distinguish from inherent measurement variability. Cold Bay and Shemya, both in Alaska and strongly influenced by marine air, are two exceptions; at these locations differences due to ocean flux are large enough to be distinguished.

### 3.2. Comparisons With HIPPO Aircraft Observations

Beginning in 2009, the High-Performance Instrumented Airborne Platform for Environmental Research Pole-to-Pole Observations (HIPPO) project has measured atmospheric CO<sub>2</sub>, other trace gases, and aerosols in a series of field campaigns. HIPPO missions consist of north-south transects spanning the Pacific Ocean from 85°N to 67°S with profiles of atmospheric trace gases every 2.2° latitude [Wofsy *et al.*, 2011]. HIPPO data have been used extensively for validation and calibration of remote sensing data products [e.g., Wunch *et al.*, 2010; Inoue *et al.*, 2013], for evaluation of atmospheric transport models [e.g., Keppel-Aleks *et al.*, 2013; Mann *et al.*, 2012], and to better understand emission and transport processes [e.g., Keppel-Aleks *et al.*, 2012; Kipling *et al.*, 2013].

Three of the five planned HIPPO missions occurred during the 2009–2010 period. Because the first deployment occurred in January 2009, shortly after the beginning of the CMS target period, our results focus on the second and third HIPPO deployments (HIPPO-2 and HIPPO-3). Both campaigns sampled the Pacific Ocean between 150°E and 90°W. During the second deployment, the majority of samples were collected between 180°W and 150°W, while during the third campaign, sampling was focused between 170°W and 140°W. HIPPO data shown are 10 s average data. GEOS-5 pseudodata are created by sampling the appropriate 3 h model output, interpolating model profiles to the measurement latitude and longitude, and then interpolating resulting profiles vertically to replicate the measurement conditions as well as possible.

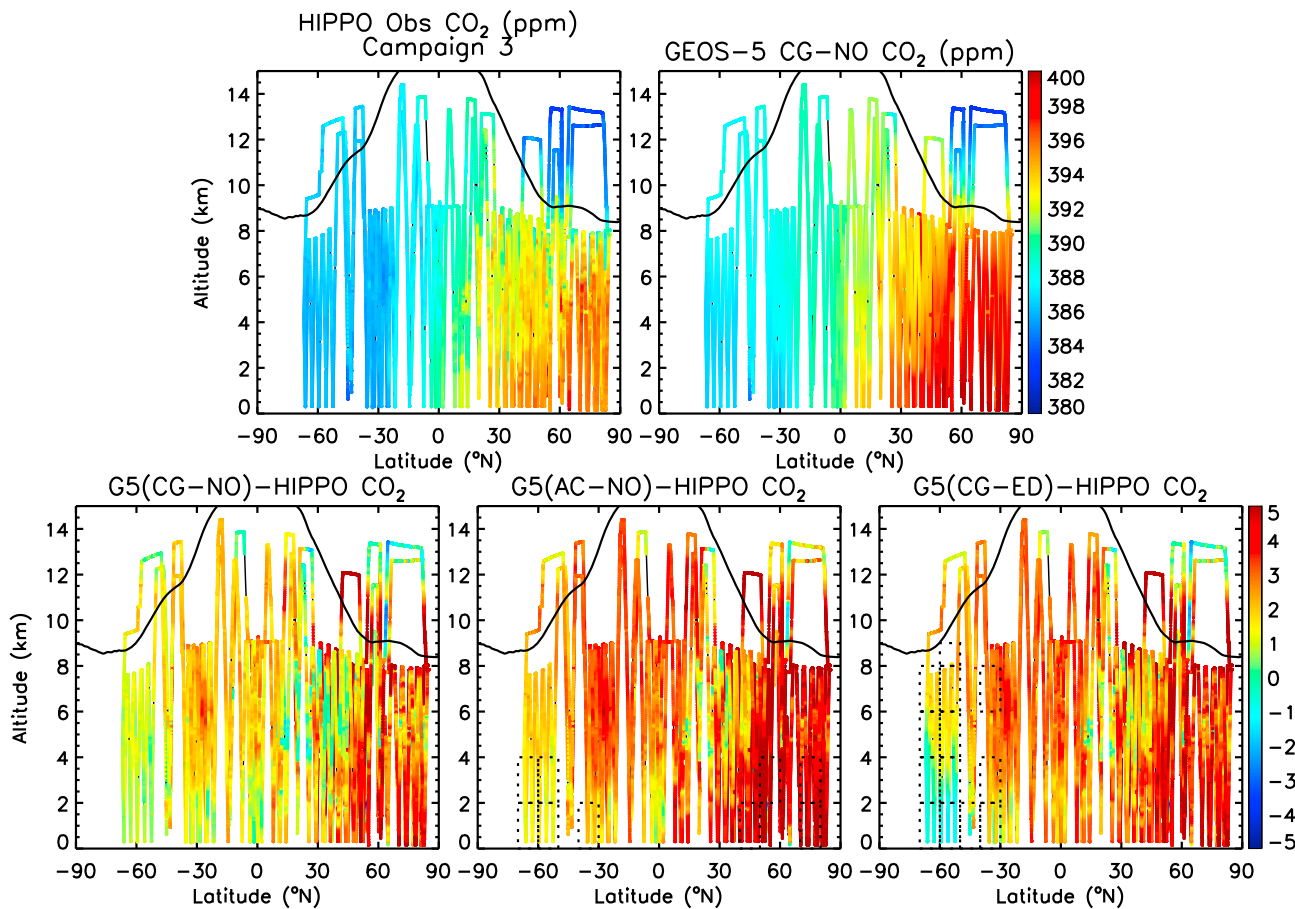




**Figure 9.** (top) Comparison of observed and simulated  $\text{CO}_2$  mixing ratios during HIPPO-2 in October and November 2009 shows HIPPO observations (top left), GEOS-5 using flux combination CG-NO (top right, no detrending) and the simulated minus observed difference (bottom left). (bottom) The difference between GEOS-5 and observations using alternate flux scenarios AC-NO (bottom middle) and CG-ED (bottom right). Dashed boxes indicate regions where flux differences are significantly larger than the variance of observations.

Figure 9 shows a longitudinal cross-section comparison between the observed  $\text{CO}_2$  data and GEOS-5 simulations (not detrended) for HIPPO-2, which occurred during October and November 2009. Relative to the HIPPO data, GEOS-5 assuming flux combination CG-NO tends to overestimate  $\text{CO}_2$  throughout the sampled region of the Pacific. Errors are greater in the NH (1–5 ppm) than the SH (less than 3 ppm). The model is able to capture the major features of the observed transect, including elevated near surface  $\text{CO}_2$  mixing ratios at NH high latitudes and enhanced  $\text{CO}_2$  mixing ratios aloft at 30°S and 30°N. Transects simulated using flux combinations AC-NO and CG-ED are generally similar to the results produced by flux combination CG-NO, but with some notable exceptions. The differences between observations and the CG-NO and AC-NO simulations show that at NH locations, overestimates in  $\text{CO}_2$  are worse when using Ames CASA fluxes compared to CASA-GFED. Near 30°S, the Ames CASA fluxes improve the model's underestimate of near surface  $\text{CO}_2$  during SH spring. However, throughout most of the campaign Ames CASA fluxes result in greater errors because the difference in the magnitude of the global land sink is less realistic. The use of ECCO2-Darwin fluxes in flux combination CG-ED results in stronger ocean drawdown at NH high latitudes during autumn. This helps to reduce the overestimate in  $\text{CO}_2$  mixing ratios at these locations, but elsewhere, model performance is degraded because of the overall weaker ocean sink.

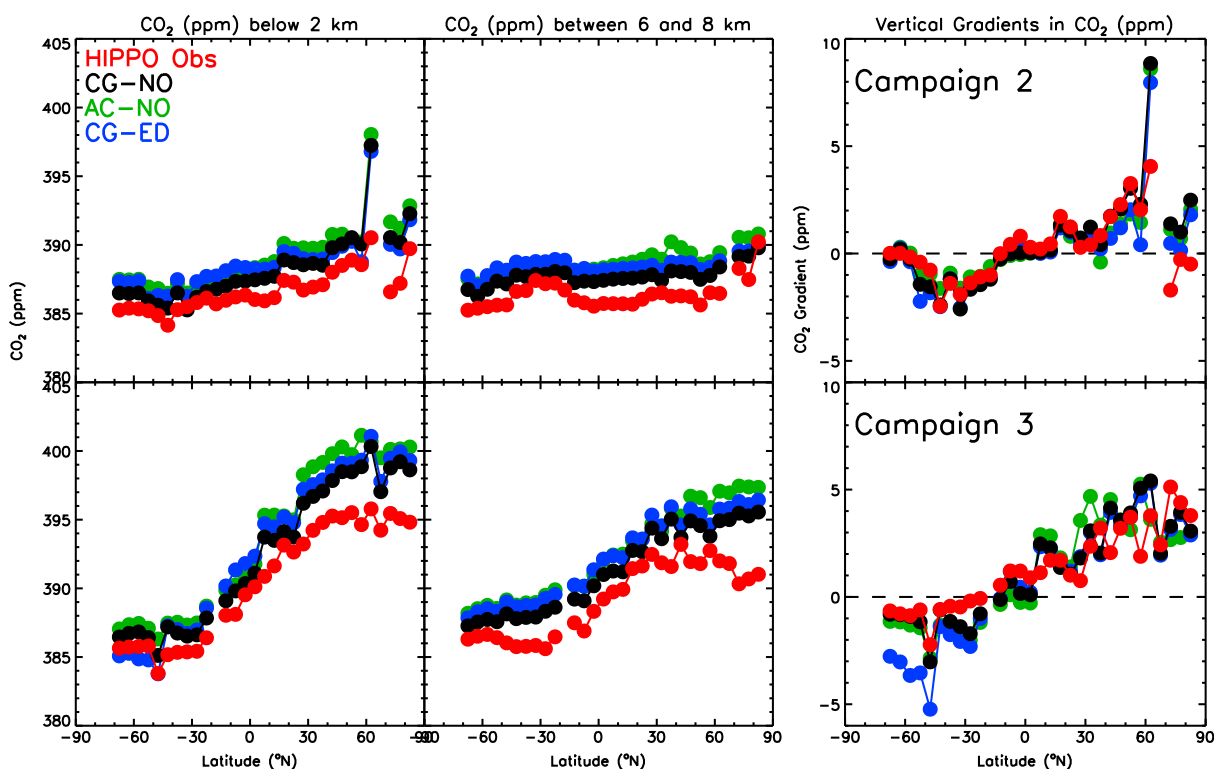
During HIPPO-3, conducted during March and April 2010 (Figure 10), model performance is generally similar to performance during the second deployment. All flux combinations result in overestimates in NH  $\text{CO}_2$ , but this is worse when using Ames CASA fluxes than when using CASA-GFED fluxes. During March and April, the much stronger Southern Ocean sink in ECCO2-Darwin is evident in observations near 60°S, resulting in larger errors in this region compared to NOBM.



**Figure 10.** (top) Comparison of observed and simulated  $\text{CO}_2$  mixing ratios during HIPPO-3 in March and April 2010 shows HIPPO observations (top left), GEOS-5 using flux combination CG-NO (top right, no detrending) and the simulated minus observed difference (bottom left). (bottom) The difference between GEOS-5 and observations using alternate flux scenarios AC-NO (bottom middle) and CG-ED (bottom right). Dashed boxes indicate regions where flux differences are significantly larger than the variance of observations.

We also evaluate the ability of HIPPO-2 and HIPPO-3 data to discern between flux combinations in different regions of the atmosphere following *Huntzinger et al.* [2011]. Observation variance is estimated here as the variance of  $\text{CO}_2$  observations within  $10^\circ$  latitude bins for layers from the surface to 2 km, from 2 to 4 km, from 4 to 6 km, from 6 to 8 km, and from 8 to 9 km. The chi-squared test statistic is calculated for each of these latitude-altitude regions using the number of observations, the mean squared difference between simulations, and the observation variance. Dashed lines in Figures 9 and 10 indicate regions where the test of variance indicates that flux difference induces differences in atmospheric  $\text{CO}_2$  mixing ratios that are statistically significant at the 95% confidence level. During HIPPO-2, land flux differences are primarily detectable in the SH midlatitudes. SH  $\text{CO}_2$  differences are smaller than those in the NH, but data exhibit much less variability (observation variances of 0.1 to 0.2 ppm in the SH compared to 0.2 to 1.4 ppm in the NH) making the differences caused by land flux easier to distinguish from the background variability of measurements. Flux differences in the SH midlatitudes also tend to be distinguishable through a deep layer of the atmosphere, not only near the surface. Ocean fluxes are difficult to distinguish during the November time frame of HIPPO-2 because the greatest disparity between ocean fluxes is in NH high latitudes where background  $\text{CO}_2$  variability is strong. Land fluxes during HIPPO-3, in contrast to HIPPO-2, are detectable only below 4 km and at midlatitude to high latitude of both hemispheres. Ocean flux differences are strongly evident in the SH high and midlatitude locations where observation variances are less than 0.1 ppm.

The HIPPO data sets are also valuable for evaluating the model's ability to simulate vertical gradients in atmospheric  $\text{CO}_2$ . Vertical gradients are calculated by first sampling model output at the measurement locations to create pseudodata. Mean values of observations and pseudodata are then calculated over

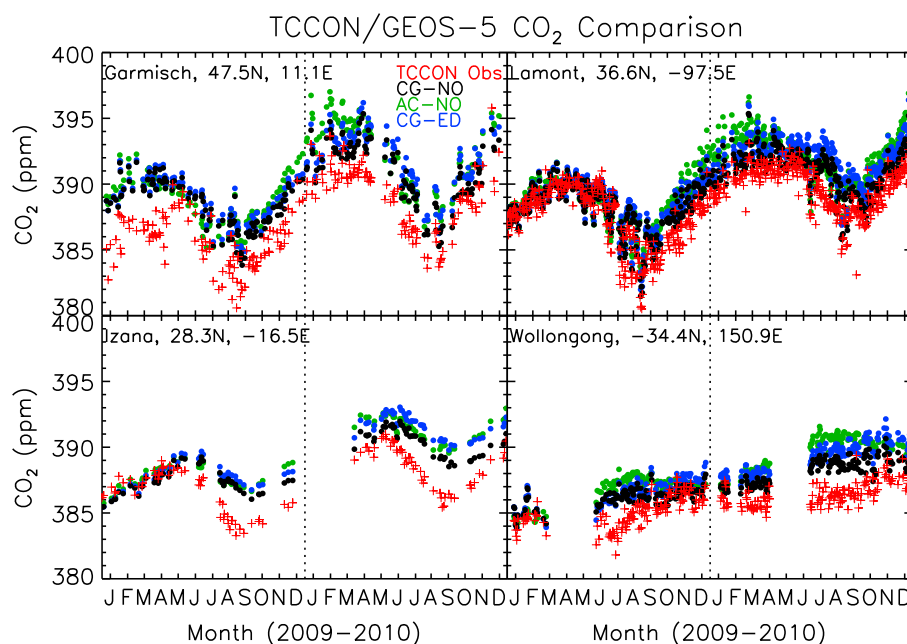


**Figure 11.** (right) Simulated and observed vertical gradients in situ  $\text{CO}_2$  derived from HIPPO data and GEOS-5 model simulations. Gradients are calculated by binning observations and pseudodata into 5 latitude bins and calculating the difference between the mean mixing ratios (left) below 2 km and (middle) between 6 and 8 km.

5° latitude bins using all data below 2 km and between 6 and 8 km in each bin. The vertical gradients for each latitude bin are calculated by subtracting the average 6–8 km  $\text{CO}_2$  mixing ratio from the 0–2 km average (Figure 11). During HIPPO-2, GEOS-5 tends to reasonably reproduce the observed interhemispheric gradient both at low levels and aloft. During HIPPO-3, all flux combinations result in the model overestimating the gradient between Northern and Southern Hemispheres  $\text{CO}_2$ . However, the vertical gradient is reasonably well simulated with the exception of ECCO2-Darwin which results in much stronger negative vertical gradients at midlatitude and high-latitude SH locations than either simulated by NOBM or observed. During HIPPO-3, the earlier spring drawdown in the NH Ames CASA fluxes also results in a slight overestimate in the vertical gradient in  $\text{CO}_2$  at most locations. The fact that the interhemispheric gradient in  $\text{CO}_2$  is overestimated during HIPPO-3, while local vertical gradients appear reasonable could indicate either errors in the assumed source distributions and magnitudes or an error in large-scale transport pathways.

### 3.3. Comparison With TCCON Column $\text{CO}_2$ Observations

The Total Carbon Column Observing Network (TCCON), established in 2004, is a network of ground-based Fourier transform spectrometers recording near-infrared direct solar spectra for a number of atmospheric trace gases including  $\text{CO}_2$  [Wunch *et al.*, 2011a]. Data are collected continuously during daylight hours when viewing is not obscured by optically thick cloud and aerosol. Observations are estimated to have a precision as high as 0.25% (~ 1 ppm) under clear-sky conditions. In contrast to the NOAA ESRL network, TCCON provides total column  $\text{CO}_2$  observations comparable to the observations made by GOSAT. TCCON is also much smaller than the NOAA flask network with 16 stations operating during the 2009–2010 study period considered here (Figure 4). TCCON observations are the primary calibration and validation data set for GOSAT [Wunch *et al.*, 2011b] and their use for flux inference has been demonstrated by Chevallier *et al.* [2011]. For this analysis, TCCON observations are slightly adjusted to account for laser sampling errors noted by Messerschmidt *et al.* [2012] and Dohe *et al.* [2013].

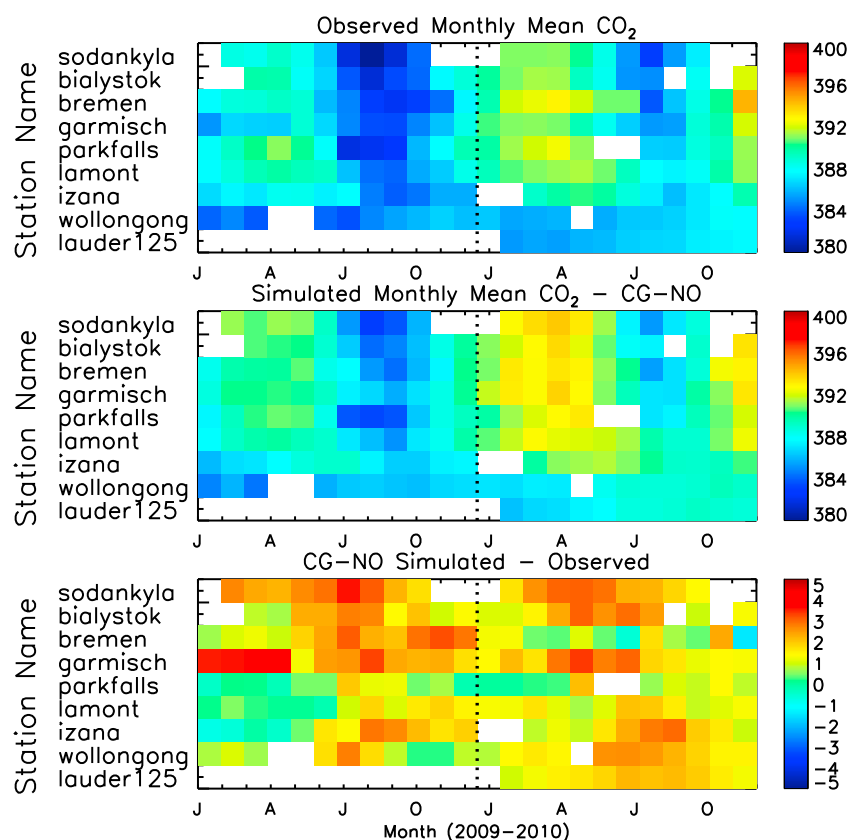


**Figure 12.** Comparison of daily mean simulated (without detrending) and observed column  $\text{CO}_2$  at Garmisch, Lamont, Izana, and Wollongong TCCON stations for 2009–2010.

In order to compare with TCCON observations, GEOS-5  $\text{CO}_2$  fields (with no detrending applied) are sampled at the appropriate observation time and convolved with the averaging kernel and a priori profile information appropriate for each station as described in Wunch *et al.* [2010]. Simulated and observed daily means of column  $\text{CO}_2$  are then calculated; the results for four stations are shown in Figure 12. As is expected, column  $\text{CO}_2$  observations exhibit smaller seasonal variations than do surface observations. At both high NH midlatitude stations shown (Garmisch and Lamont), GEOS-5 simulations are able to reasonably reproduce the observed column  $\text{CO}_2$  mixing ratios with a slight underestimation in the amplitude of the seasonal cycle (average underestimate over all NH midlatitude and high-latitude stations of 1.4 ppm for CG-NO fluxes). Garmisch observations indicate a small decrease in column  $\text{CO}_2$  from January to March 2009 which is not reproduced by any model configuration; a similar feature is not present at other NH TCCON stations or at NOAA surface locations. Differences due to the flux combination assumed are small, typically no more than a few parts per million, but the difference in seasonal cycle between CASA-GFED3 and Ames CASA is evident particularly during NH spring. At the Izana, Tenerife station, GEOS-5 simulations overestimate summer  $\text{CO}_2$  mixing ratios relative to observations, a feature also found in the comparison against Mauna Loa surface observations (Figure 5).

Monthly mean errors between GEOS-5 simulations (no detrending applied) and TCCON observations (Figure 13) show that model errors are typically less than 3 ppm when flux combination CG-NO is assumed. The model tends to overestimate  $\text{CO}_2$  at NH locations during spring and summer, while during winter months, errors are small (less than 1 ppm at most stations). Model errors at NOAA ESRL NH stations (Figure 6) during winter are greater than 5 ppm at some locations, resulting in small overestimates in seasonal cycle amplitude at the surface (Figure 7). In contrast, the model slightly underestimates the seasonal cycle amplitude at NH TCCON locations. This may be because of errors in simulated transport during winter months resulting in too much  $\text{CO}_2$  near the surface, or because of differences in sampling locations between the ESRL and TCCON networks.

Figure 14 shows differences in column  $\text{CO}_2$  mixing ratio at TCCON stations due to the difference in land and ocean fluxes assumed (no detrending applied). The seasonality of differences due to differences between Ames CASA and CASA-GFED3 is similar to the analysis of surface observations, but the magnitude of flux differences manifests as a smaller magnitude atmospheric mixing ratio difference when the column is considered. Despite the smaller total magnitude of the column differences, they are distinguishable from



**Figure 13.** (top) Monthly mean CO<sub>2</sub> (ppm) observed at TCCON stations during 2009–2010 compared with simulated CO<sub>2</sub> assuming flux combination CG-NO (middle, no detrending). (bottom) The model-observation difference. Stations are oriented from north (top) to south (bottom) on all plots.

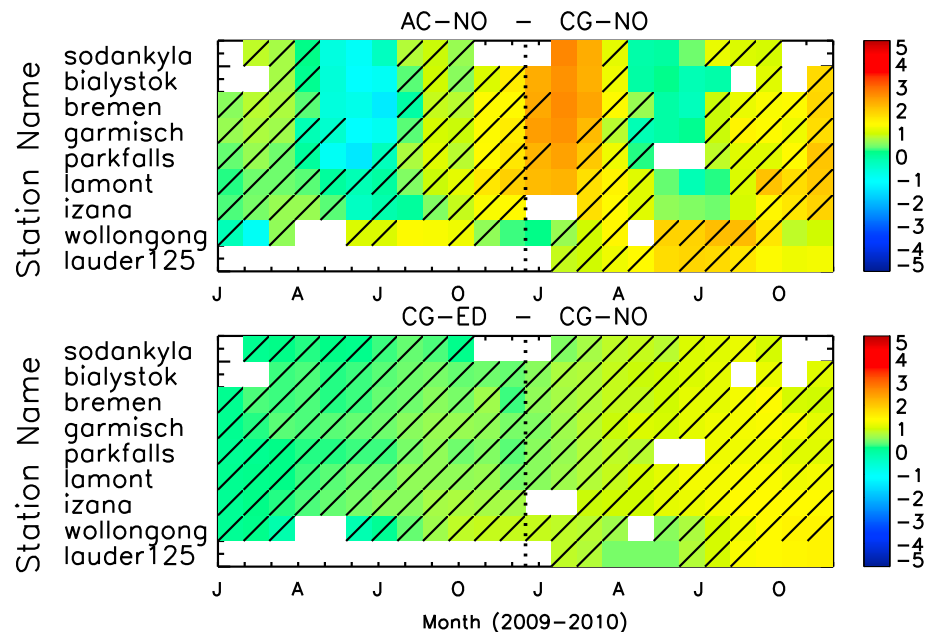
inherent measurement variability at most NH stations during late spring and early summer, and winter 2010. This is largely due to the weaker variability in the column data and greater data volume yielded by continuous measurements. The Ames CASA fluxes result in lower CO<sub>2</sub> (0.5–1.5 ppm) at NH stations during MAM 2009, and higher (2–3 ppm) CO<sub>2</sub> at NH midlatitude and high-latitude locations during 2010 winter, degrading performance relative to the CASA-GFED3 fluxes.

Detecting ocean flux differences at TCCON locations is considerably more difficult than detecting land flux differences. Data from the Lauder station show some ability to differentiate between ocean fluxes during some months of 2010, but data are not available at this site during 2009. These results point to the difficulty of detecting ocean CO<sub>2</sub> flux differences with column observations. The nearest surface station, Baring Head, New Zealand, indicates a higher percentage of data useful for differentiating between ocean fluxes throughout the year. Atmospheric mixing ratio differences caused by ocean flux uncertainty are typically smaller than differences caused by land flux and, as a result, more difficult to separate from inherent measurement variability. Additionally, the TCCON network includes no stations further south than Lauder, making it impossible to directly observe Southern Ocean flux differences.

### 3.4. Comparisons With AIRS Satellite Observations

The Atmospheric Infrared Sounder (AIRS) instrument aboard the Aqua satellite provides a record of midtropospheric CO<sub>2</sub> from 2002 to present. Because data are collected during both day and night and retrievals are performed in the presence of clouds, AIRS data provide unprecedented global coverage. Accuracies are reported to be 2 ppm or better by *Chahine et al.* [2005] with a nadir footprint of 90 by 90 km<sup>2</sup> [*Chahine et al.*, 2008]. AIRS CO<sub>2</sub> observations have been used to study temporal [e.g., *Jiang et al.*, 2010; *Li et al.*, 2010] and spatial [e.g., *Ruzmaikin et al.*, 2012; *Bai et al.*, 2010] variability in midtropospheric CO<sub>2</sub>, to evaluate





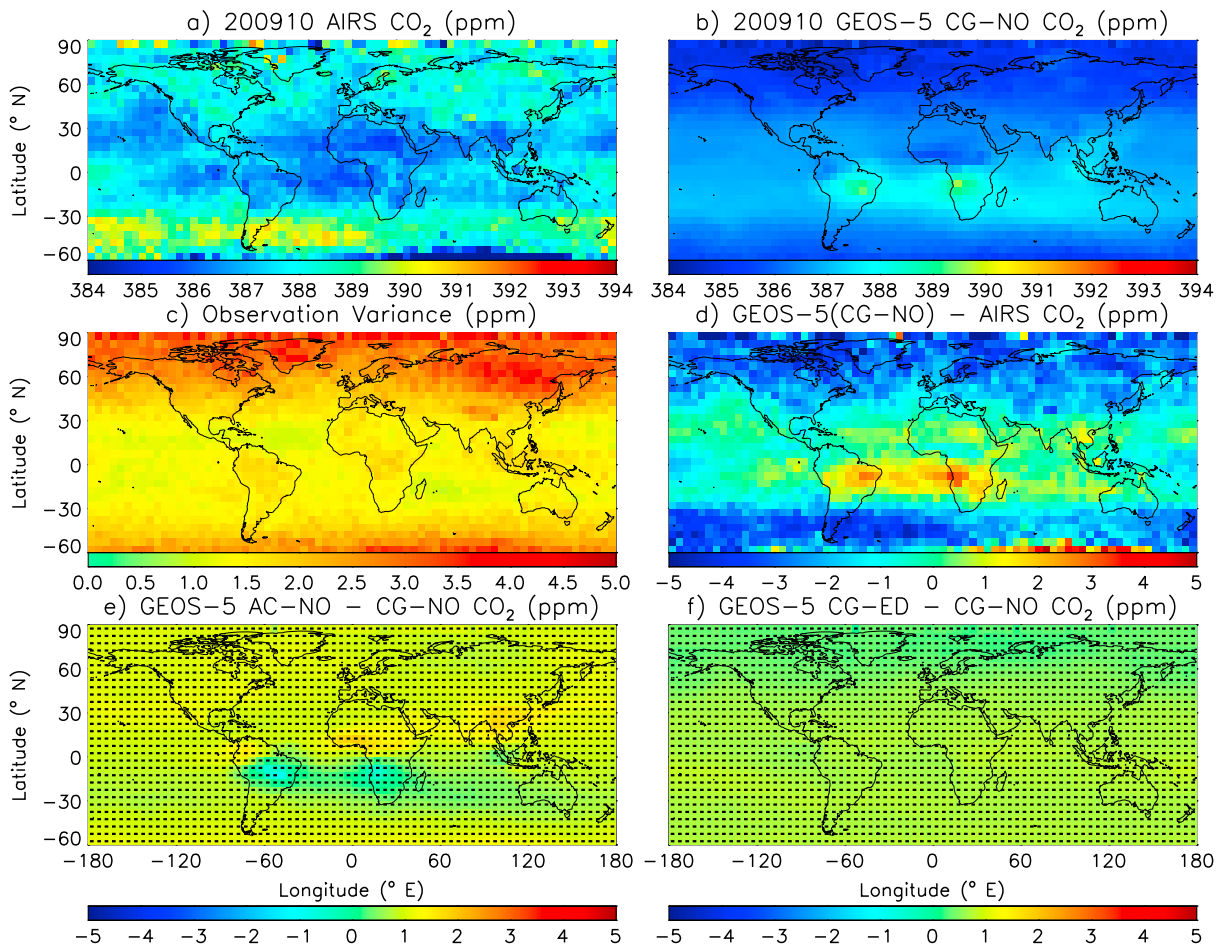
**Figure 14.** Differences in monthly mean CO<sub>2</sub> mixing ratio (no detrending) at TCCON stations due to (top) land flux estimates and (bottom) ocean flux estimates. Diagonal bars indicate instances in which flux differences would not be statistically significant from observation variability.

atmospheric CO<sub>2</sub> simulations [e.g., Feng et al., 2011], and for data assimilation studies [e.g., Liu et al., 2012; Engelen et al., 2009].

GEOS-5 fields (without detrending) are sampled in the same locations as AIRS version 5 CO<sub>2</sub> observations using level 2 swath data. Simulated CO<sub>2</sub> profiles for the appropriate 3 h period are interpolated to the location of each observation that passes AIRS quality assurance procedures. AIRS CO<sub>2</sub> weighting functions specific to the observation latitude and background CO<sub>2</sub> mixing ratio are applied as described in Chahine et al. [2008]. Monthly mean simulated and observed CO<sub>2</sub> mixing ratios are then calculated over 5° latitude by 5° longitude regions.

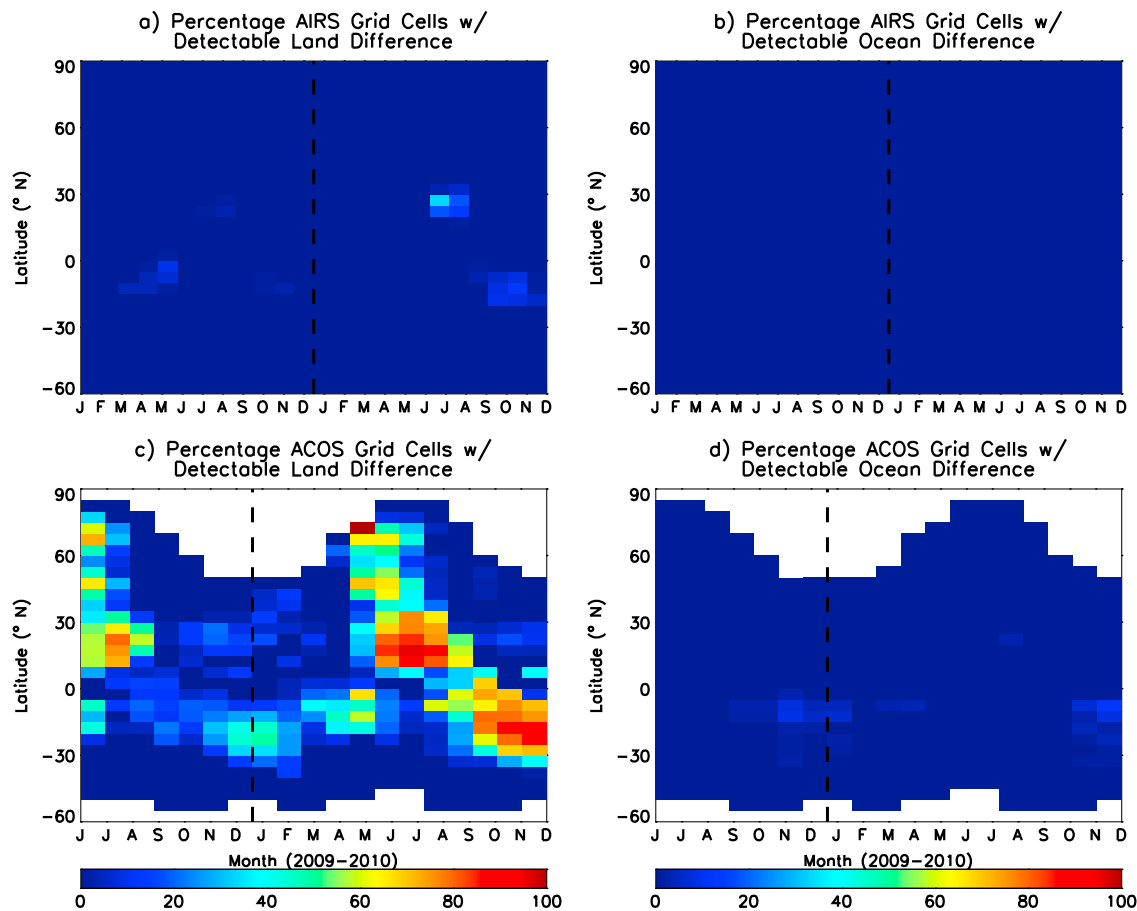
Figure 15 presents a comparison between AIRS observations and GEOS-5 simulated CO<sub>2</sub> for October 2009, a month that coincides with the beginning of the HIPPO-2 campaign. There is little agreement between the AIRS observations and the CG-NO GEOS-5 simulation. AIRS indicates peak CO<sub>2</sub> mixing ratios in SH midlatitudes with minimum values in tropics and subtropics of both hemispheres. In contrast, GEOS-5 indicates enhanced midtropospheric CO<sub>2</sub> over South America and SH Africa, likely because of the lofting of fire emissions in these regions (Figure 2). GEOS-5 does not indicate the presence of enhanced CO<sub>2</sub> mixing ratios in the midlatitudes of either hemisphere resulting in a 2–5 ppm underestimate in these regions relative to AIRS. In biomass burning regions, however, GEOS-5 overestimates midtropospheric CO<sub>2</sub> by 1–3 ppm relative to AIRS. These results are in sharp contrast to the comparison between HIPPO-2 CO<sub>2</sub> observations and GEOS-5 presented in Figure 9; that comparison indicates that GEOS-5 tends to overestimate midtropospheric CO<sub>2</sub> in the 30–60°S band over the Pacific Ocean from 0 to 8 km. The discrepancy between AIRS and GEOS-5 in this region cannot be easily explained by errors in GEOS-5 vertical transport because Figure 11 indicates that GEOS-5 vertical CO<sub>2</sub> gradients throughout the SH differ from observed vertical gradients by less than 1 ppm. NH underestimates of GEOS-5 relative to AIRS also coincide with regions where HIPPO-2 data indicate that GEOS-5 is overestimating CO<sub>2</sub> by 2–4 ppm in the midtroposphere.

AIRS data show a significant amount of temporal variability evident in the observation variance calculated for each 5° by 5° grid box (Figure 15). Observation variance is calculated for each grid cell by computing daily means of available data during the 2009–2010 period, fitting a smooth curve through the daily data, and calculating the standard deviation of the residuals, similar to the calculation of observation variance for NOAA ESRL and TCCON data. Using this method, observation variance is typically between 1.5 and 4 ppm



**Figure 15.** Comparison of AIRS observed midtroposphere CO<sub>2</sub> mixing ratio with GEOS-5 simulations for October 2009 for 5° latitude by 5° longitude grid boxes. (a) AIRS monthly mean mixing ratio and (b) monthly mean GEOS-5 mixing ratios (no detrending) assuming flux combination CG-NO sampled using AIRS pressure weighting functions. (c) The variance of AIRS observations for each grid box and (d) GEOS-5 CO<sub>2</sub> minus AIRS. The difference between (e) GEOS-5 simulated CO<sub>2</sub> for flux combinations AC-NO and CG-NO and (f) combinations CG-ED and CG-NO; grid cells where these differences are not statistically significant at the 95% confidence level are indicated by black lines.

with the largest observed variability in the midlatitude and high latitude of both hemispheres. Differences between GEOS-5 simulations assuming different flux combinations are typically small in the midtroposphere and upper troposphere regions represented by AIRS observations. The background difference between flux combinations CG-NO and AC-NO, due to the difference in annual growth rate, is 1 ppm during October. In biomass burning regions in SH Africa and South America, Ames CASA fluxes result in CO<sub>2</sub> mixing ratios 1 ppm less than CASA-GFED fluxes, or nearly 2 ppm if the difference in annual growth rate were removed. CASA-GFED results in 2 ppm more CO<sub>2</sub> over equatorial Africa and SE Asia, but if the difference in growth rate is removed, these differences are only 1 ppm. In tropical regions, where vertical transport of surface fluxes is most rapidly communicated to the midtroposphere and upper troposphere, spatial patterns of atmospheric mixing ratio differences resemble underlying surface flux differences shown in Figure 2. However, the magnitude of these differences (1–2 ppm) is typically smaller than the observation variance leading to a lack of ability to discern between flux scenarios with any statistical significance. Ocean flux differences (shown in Figure 15f), which are smaller in magnitude but more consistent throughout the year, are not evident when model simulations are sampled with AIRS pressure weighting functions except as a very slight modification to the north-south CO<sub>2</sub> gradient that is not statistically significant. AIRS is not able to observe the large ocean flux differences in high-latitude regions; data are not available south of 60°S and in NH high latitudes, slow vertical transport times produce no spatial signature of the underlying flux differences in the midtroposphere and upper troposphere where AIRS measurements are most sensitive.



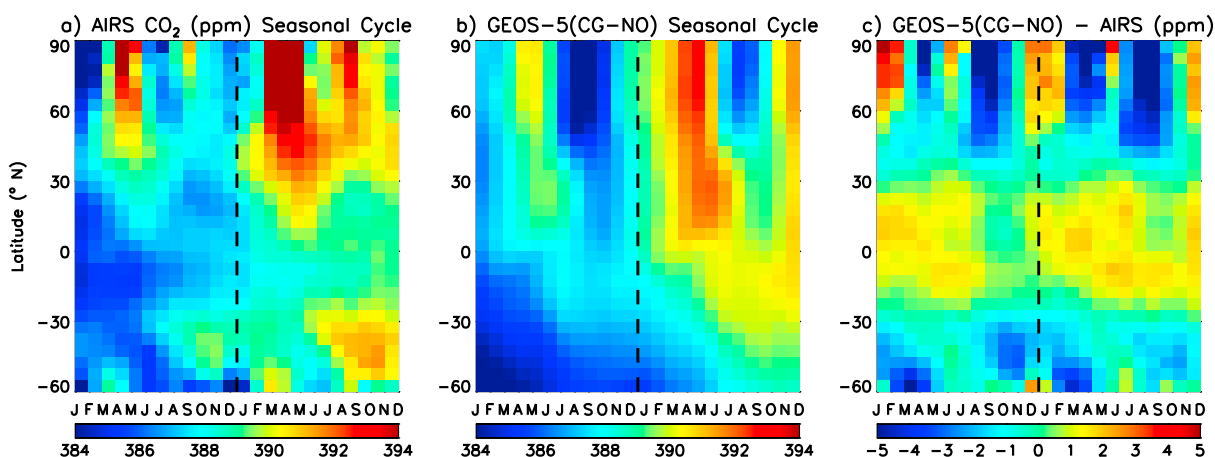
**Figure 16.** Percentage of 5° latitude by 5° longitude grid cells per month and per latitude bin in which (a and c) land (b and d) ocean flux differences manifest as significantly different satellite CO<sub>2</sub> mixing ratios. Figures 16a and 16b show the detectability of flux differences by AIRS, while Figures 16c and 16d show results for ACOS GOSAT data.

In all months during 2009–2010, AIRS data would have little utility for differentiating between spatial patterns of flux differences. Figures 16a and 16b show the percentage of 5° grid boxes per latitude bin that contain statistically significant land and ocean flux differences for each month during the study period. Land flux differences would seldom be detectable and only in a small number of grid cells in the tropics and subtropics. Ocean flux differences are not detectable in any month or location because the magnitude of the mixing ratio differences is always smaller than the variability inherent in the measurements.

Figure 17 shows a comparison between simulated and observed monthly zonal mean mixing ratios derived from AIRS data and model pseudodata. GEOS-5 does not capture the seasonal cycle of zonal mean CO<sub>2</sub> observed by AIRS. GEOS-5 midtropospheric CO<sub>2</sub> exhibits a much stronger minimum during the NH growing season (July–October) than does AIRS. AIRS indicates much stronger maxima in the NH high latitudes during spring than indicated by GEOS-5, and a secondary maxima during September that is not simulated by GEOS-5. GEOS-5 tends to produce larger than observed CO<sub>2</sub> mixing ratios in the tropics and subtropics throughout the year. In the SH midlatitudes, GEOS-5 underestimates AIRS zonal mean CO<sub>2</sub> with the largest differences in spring and autumn.

### 3.5. Comparisons With GOSAT Satellite Observations

The GOSAT, launched in 2009 by the Japanese Aerospace Exploration Agency (JAXA), monitors CO<sub>2</sub> and methane from space by analyzing high-resolution spectra of reflected sunlight within several near-infrared bands [Kuze *et al.*, 2009; Yokota *et al.*, 2009]. The near-infrared measurement technique allows for greater sensitivity to CO<sub>2</sub> near the surface than the thermal infrared measurements of AIRS that are sensitive



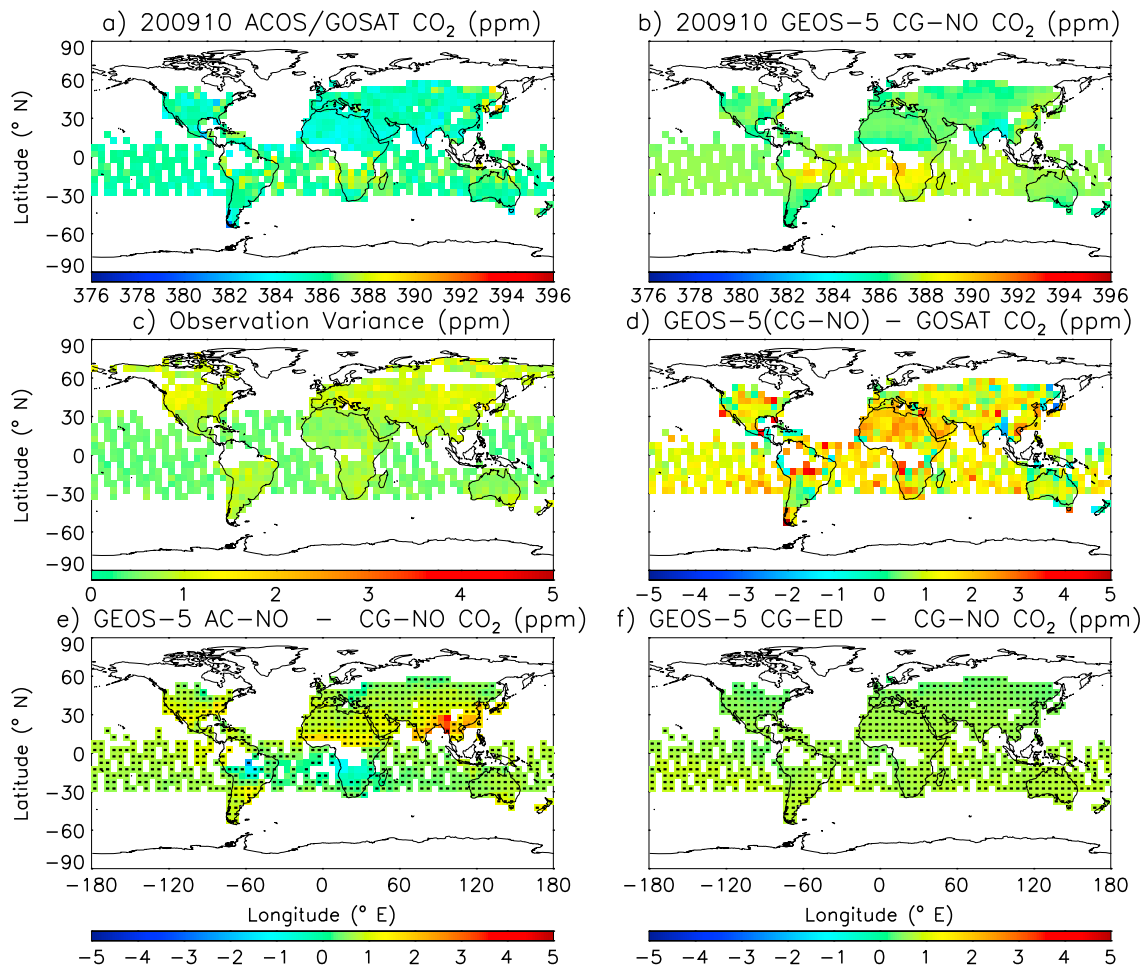
**Figure 17.** Seasonal cycle of (a) zonal mean AIRS observed mixing ratios, (b) zonal mean GEOS-5 (CG-NO) mixing ratios (no detrending applied) calculated with AIRS weighting functions, and (c) the difference between zonal mean model results and observations. All units are in parts per million.

primarily to the midtroposphere and upper troposphere. Retrievals are only performed under cloud-free conditions resulting in a much smaller yield of usable soundings than AIRS. GOSAT observations have been used to examine  $\text{CO}_2$  mixing ratios in megacities [Kort *et al.*, 2012] and to estimate regional carbon fluxes [e.g., Takagi *et al.*, 2011; Basu *et al.*, 2013; Maksyutov *et al.*, 2013]. Several different retrievals of column averaged  $\text{CO}_2$  dry air mole fraction ( $\text{XCO}_2$ ) based on GOSAT observations are currently being produced. In this work, we use data produced by NASA's Atmospheric  $\text{CO}_2$  Observations from Space (ACOS) effort, a collaboration between the original JAXA GOSAT team and NASA's Orbiting Carbon Observatory (OCO) science team [Crisp *et al.*, 2012]. These data are produced using a modified version of the OCO retrieval algorithm applied to GOSAT observed spectra as described in O'Dell *et al.* [2012] and have been validated extensively using TCCON measurements [e.g., Wunch *et al.*, 2010, 2011a, 2011b]. Based on retrievals of realistic simulated observations, O'Dell *et al.* [2012] estimate that observations contain root-mean-square  $\text{XCO}_2$  errors of  $\sim 1$  ppm and a positive bias of 0.3 ppm. Dates used here are ACOS version 3.4.

GEOS-5 fields (without detrending applied) are sampled at GOSAT measurement times and locations. Model profiles are interpolated to the 20 atmospheric pressure levels used in the retrieval process and averaging kernels provided as part of the ACOS data product are applied to convolve GEOS-5 mixing ratios with the a priori  $\text{CO}_2$  mixing ratio profile.  $\text{XCO}_2$  is calculated from this simulated pseudo data profile. Monthly mean simulated and observed  $\text{CO}_2$  mixing ratios are then calculated over  $5^\circ$  latitude by  $5^\circ$  longitude regions as was done for the AIRS data comparison.

During October 2009, the spatial distribution of GEOS-5 assuming flux CG-NO compares much more favorably with GOSAT observations (Figure 18) than with AIRS (Figure 15). The distribution of GOSAT  $\text{XCO}_2$  shows fairly uniform mixing ratios globally with slightly larger  $\text{XCO}_2$  over South America and Africa. The GEOS-5 distribution is similar though  $\text{XCO}_2$  is overestimated by 1–3 ppm in most locations. This is consistent with HIPPO-2 aircraft observations (Figure 9) which show similar overestimates between 0 and 8 km. Observation variance was calculated as described above for AIRS data; observations were binned into  $5^\circ$  grid cells, daily means calculated and then the standard deviation of residuals around a smooth curve considered the observation variance. Using this technique, estimated observation variance is typically between 0.5 and 1 ppm. However, despite the smaller observation variance values, differentiating between land and ocean fluxes remains difficult because of the small number of observations collected. The difference in  $\text{XCO}_2$  due to the underlying difference in land flux would be primarily detectable over Africa. Ocean flux differences are too small to be detectable at any location.

A comparison between GOSAT data and GEOS-5 simulations during July 2009 (Figure 19) shows similar features in both models and observations. The model is able to reasonably reproduce the observed spatial distribution of  $\text{XCO}_2$  but tends to overestimate  $\text{XCO}_2$  by up to 5 ppm in most midlatitude and high-latitude locations. This is consistent with model overestimates of 3–5 ppm at surface (Figure 6) and TCCON stations



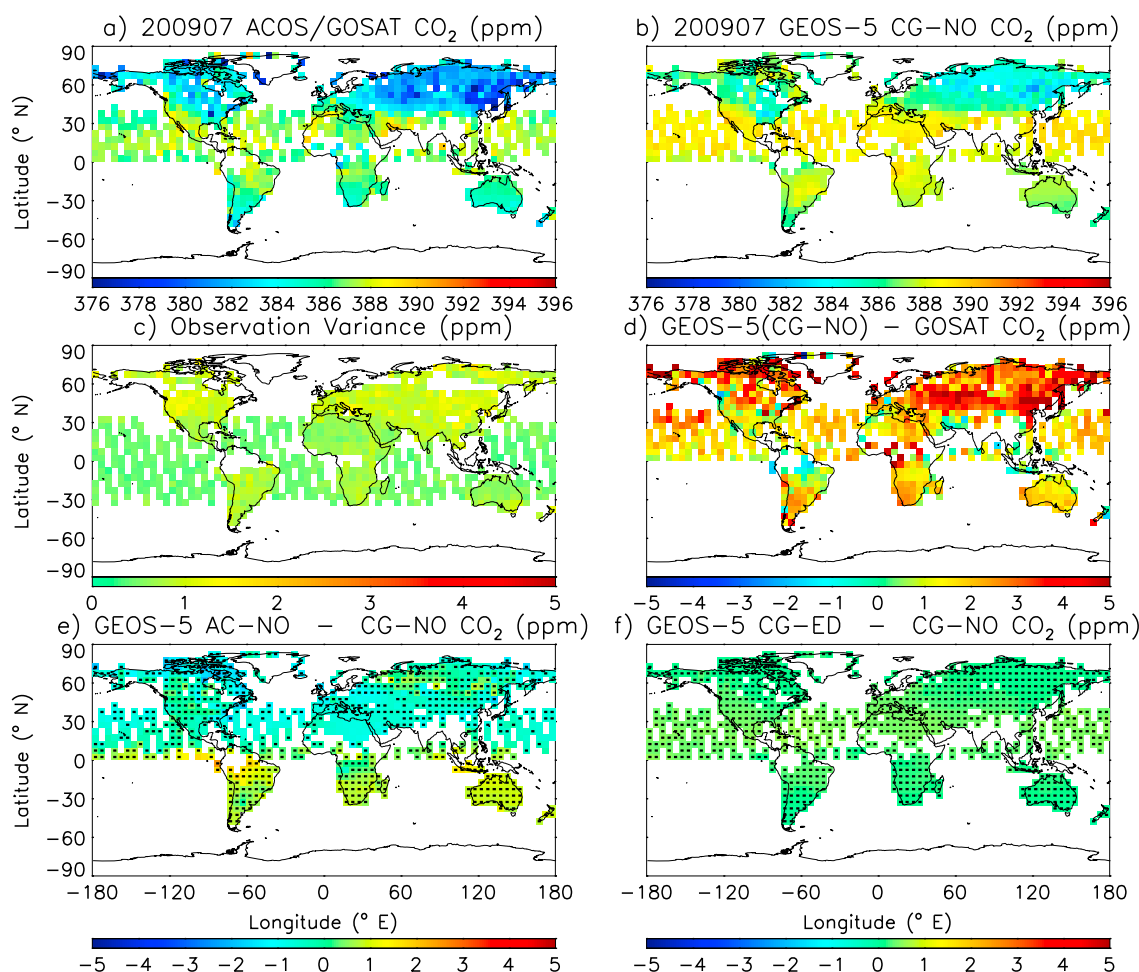
**Figure 18.** Same as Figure 14 but using ACOS retrievals of GOSAT observations and sampling instead of AIRS.

(Figure 13) during NH summer months. Model overestimates are larger during July than during October because the land sink, at its peak during NH summer, is likely weaker in both CASA-GFED and Ames CASA than in reality. The Ames CASA land fluxes, characterized by earlier drawdown of  $\text{CO}_2$  in the NH spring, result in smaller  $\text{CO}_2$  mixing ratios during July in the NH than CASA-GFED despite the fact that the annual total Ames CASA land sink is weaker. Though the largest differences in land flux during July are in the NH high latitudes (Figure 3), the resultant atmospheric mixing ratio difference is not generally detectable at these locations. Instead, land flux differences are most readily observed over NH ocean locations and over North Africa because the observations exhibit much less variability in these locations. As in October, ocean flux differences are not large enough to be detectable at any locations.

Figures 16c and 16d also show the percentage of  $5^\circ$  grid boxes per latitude bin that contain statistically significant ( $p < 0.05$ ) land and ocean flux differences for each month when GOSAT data are considered. In contrast to AIRS and in spite of very sparse sampling, GOSAT observations would be able to distinguish spatial pattern differences in atmospheric  $\text{CO}_2$  due to differing land fluxes in both hemispheres primarily during their spring and summer months. Ocean flux differences are statistically significant in only a small number of grid cells in the SH tropics.

Simulated zonal mean  $\text{XCO}_2$  from GEOS-5 agrees much more favorably with GOSAT data (Figure 20) than with AIRS (Figure 17) as was the case when spatial distributions of  $\text{CO}_2$  during October 2009 were considered. GEOS-5 tends to overestimate  $\text{XCO}_2$  throughout the year with the largest degree of overestimation during NH spring and summer. The use of Ames CASA instead of CASA-GFED land fluxes helps to improve the



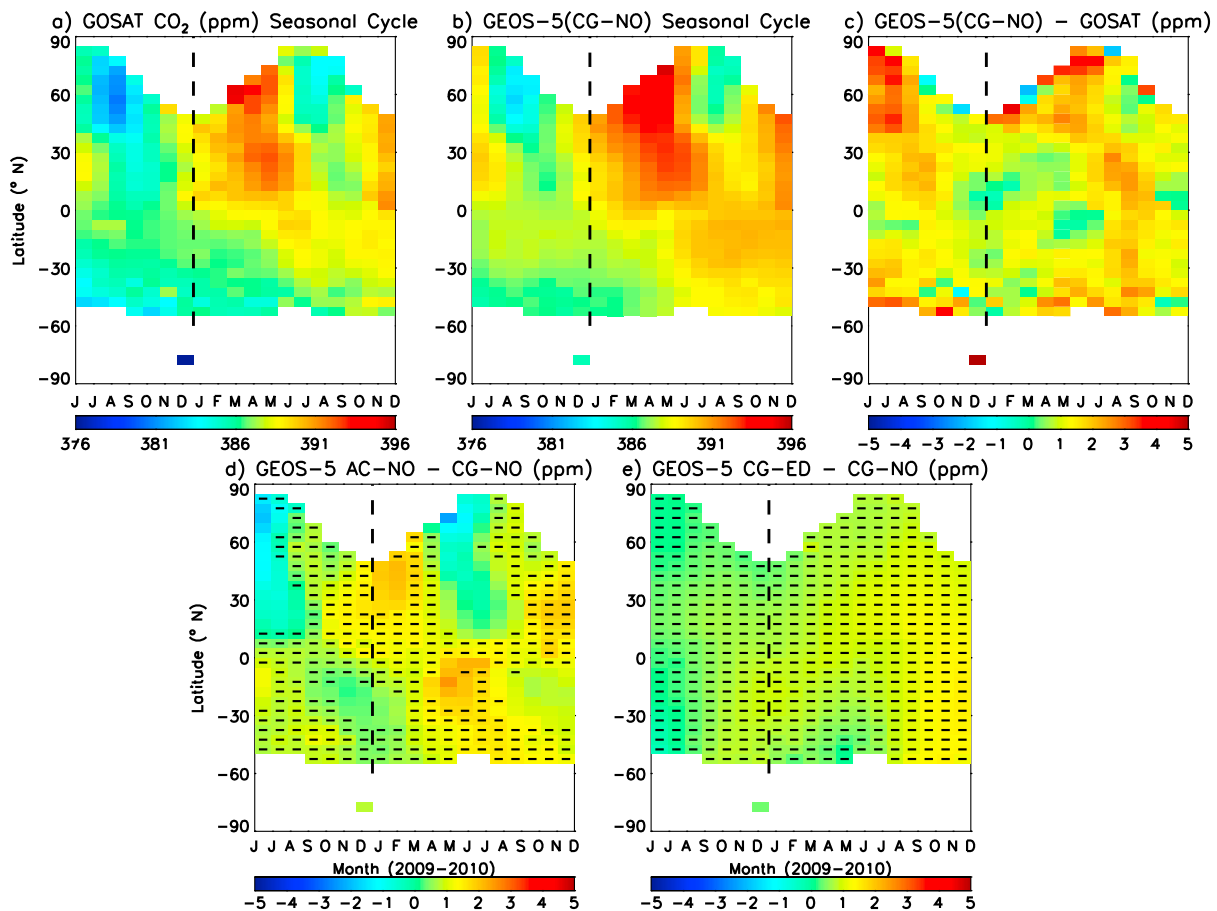


**Figure 19.** Same as Figure 17 but for July 2009.

overestimation during spring but degrades the comparison with observations during NH winter. Atmospheric CO<sub>2</sub> differences caused by differing land fluxes are detectable when zonal means are considered primarily during spring and summer months in both hemispheres. Ocean flux differences are too small to be detected by GOSAT observations in all months and in all locations.

#### 4. Summary and Conclusions

As part of NASA's CMS Flux Pilot Project, the GEOS-5 AGCM was used to simulate atmospheric CO<sub>2</sub> mixing ratios during 2009 and 2010 using two sets of model-based land and ocean flux estimates. All land and ocean flux estimates were informed by multiple satellite data sets and compared with numerous observations with the goal of better understanding the constraint on carbon flux provided by current observing systems. Despite incorporating information from multiple satellites, the two land models used in this work differed by an average of 1.7 Pg C in their estimate of the global land carbon sink during 2009 and 2010. Both land flux models estimate the land biosphere as a net source of carbon to the atmosphere rather than a ~3 Pg C sink as indicated by the Global Carbon Budget [Le Quéré *et al.*, 2013]. While the presence of a missing land sink was indicated by TRANSCOM models over 10 years ago [e.g., Gurney *et al.*, 2002] and much progress has since been made in understanding the processes which may account for the disparity between forward and inverse model flux estimates, the CMS FPP results show that estimating realistic net land carbon fluxes at both global and regional scales remains a challenge for models, even those informed by remote sensing observations. Ocean models differ by 1.4 Pg C per year in global carbon flux. In ocean waters south of 30°S, model flux differences are 25% as large as the mean regional flux and the phasing of the seasonal cycle also



**Figure 20.** Seasonal cycle of (a) zonal mean ACOS observed mixing ratios, (b) zonal mean GEOS-5 (CG-NO) mixing ratios (no detrending) calculated with ACOS weighting functions, and (c) the difference between zonal mean model results and observations. The difference in zonal mean caused by (d) land and (e) ocean flux differences. Dashed lines indicate months where flux differences are not statistically significant from inherent measurement variability. All units are in parts per million.

differs substantially. The magnitude of these differences underscores the continued uncertainty surrounding the ability of natural carbon reservoirs to compensate for increasing fossil fuel emissions.

Despite considerable differences in global flux, the performance of GEOS-5 simulations is generally similar because of the strong role played by meteorological transport. This finding is consistent with previous analyses by *Keppel-Aleks et al.* [2012]. GEOS-5 simulations assuming all flux combinations tended to overestimate surface CO<sub>2</sub> mixing ratios in the NH, particularly during spring and summer seasons when errors are 3–5 ppm. The amplitude and phasing of the seasonal cycle at most NH locations are reasonably well simulated though the model tends to slightly overestimate the amplitude of the seasonal cycle at the surface. All flux combinations examined in this study struggle to adequately reproduce the timing of the observed seasonal cycle at SH stations and at midlatitude and high-latitude SH locations and strongly overestimate the magnitude of seasonal variability by as much as 5 ppm in some locations. Because the seasonal cycle is much smaller at SH midlatitude and high latitude compared to NH locations, errors of this magnitude are greater than the observed seasonal cycle amplitude.

Comparisons with HIPPO observations provide the opportunity to evaluate the realism of simulated vertical mixing processes. As expected from flux and surface CO<sub>2</sub> comparisons, GEOS-5 tends to overestimate CO<sub>2</sub> mixing ratios relative to aircraft observations, particularly in the NH midlatitude and high latitude. Differences between mean near surface CO<sub>2</sub> and CO<sub>2</sub> observed between 6 and 8 km show that the model succeeds in simulating realistic vertical gradients in CO<sub>2</sub> during October–November 2009 and March–April 2010. The HIPPO data comparisons indicate that during these months and over the Pacific, model errors in surface and column CO<sub>2</sub> are not likely to be attributable to vertical transport errors.

GEOS-5 simulations succeed in capturing the major features of column  $\text{CO}_2$  observations from TCCON and GOSAT though the model tends to overestimate  $\text{CO}_2$  by 3–4 ppm during NH summer regardless of the combination of fluxes used. GEOS-5 is also able to reproduce the spatial patterns of  $\text{XCO}_2$  observed by GOSAT. Comparisons between GEOS-5 and surface, aircraft, TCCON, and GOSAT data reveal a consistent picture of the capabilities of contemporary models to reproduce atmospheric  $\text{CO}_2$  observations. Models tend to overestimate  $\text{CO}_2$  because the combined land and ocean flux assumed is too weak. Overestimates at the surface are largest during the NH summer months, consistent with an underestimate in the strength of land biosphere sink that is most active in this region and during this season. HIPPO observations show that vertical gradients are reasonably well simulated, meaning that excess  $\text{CO}_2$  in the model atmosphere is distributed through the deep layer (typically between the surface and 8 km) sampled by the aircraft. Total column  $\text{CO}_2$  measurements (TCCON and GOSAT) are overestimated by GEOS-5 in the same locations and seasons as indicated by comparisons with in situ measurements, but by a slightly smaller amount.

It is more difficult to reconcile AIRS midtropospheric  $\text{CO}_2$  observations with other observational records and GEOS-5. When compared with AIRS midtropospheric  $\text{CO}_2$  mixing ratios, the model is unable to represent either observed spatial distributions or seasonal cycle information. This discrepancy is noteworthy because the variability of midtropospheric partial columns would be expected to be smaller than the variability inherent in the total column measurements. It is difficult to explain the differences between GEOS-5 and AIRS as simply resulting from model errors, because analysis using aircraft data indicates that vertical transport in the model is reasonably well simulated, and because the model is able to reasonably simulate the total column observations of GOSAT and TCCON. While AIRS vertical weighting functions peak in the midtroposphere, they are strongly sensitive to the upper troposphere/lower stratosphere (UTLS). Depending on the latitude of the observation, 60–71% of the observed signal comes from pressures lower than 500 hPa. In order for GEOS-5 simulations to both agree with TCCON and GOSAT and disagree with AIRS  $\text{CO}_2$  observations, the model would have to have large errors in the UTLS including overestimates in the tropics and alternating underestimates and overestimates in the NH high latitudes. While there is no evidence of this type of model bias in the comparisons with HIPPO presented here, sampling above 8 km is fairly sparse. Further evaluation of AIRS  $\text{CO}_2$  mixing ratios is needed to evaluate the quality of reported data and uncertainty estimates as well as averaging kernels, particularly in NH high and SH midlatitudes, where features exist that appear to be inconsistent with other data sets. This work is especially important if AIRS and other  $\text{CO}_2$  observations are to be used together in joint assimilation and inversion frameworks.

Despite differences between land and ocean flux estimates greater than 1 Pg C, resulting differences in atmospheric mixing ratio at remote surface sites are small, typically less than 5 ppm at the surface and 3 ppm in the column, and difficult to distinguish from inherent observed variability. At NOAA ESRL surface stations, the difference between the Ames CASA and CASA-GFED land fluxes manifests most clearly in the atmosphere during NH spring and winter months. Outside of NH high-latitude locations,  $\text{CO}_2$  mixing ratio differences are rarely statistically significant when the variability of observations is considered. TCCON observations succeed in detecting seasonal differences between land flux estimates in NH winter and spring months. Though flux differences result in smaller mixing ratio differences in the column, the continuous monitoring strategy of TCCON produces larger quantities of data, which facilitates separation of flux differences from inherent measurement variability. AIRS satellite observations are unable to discern between land flux models because the variability of observations is larger than the magnitude of the resulting simulated mixing ratio differences in the midtroposphere and upper troposphere. GOSAT observations are better suited to observing flux differences because of greater sensitivity near the surface, and they show an ability to distinguish between land flux estimates during summer months. However, GOSAT's ability to differentiate between land flux estimates remains limited by sparse coverage.

Ocean flux differences, characterized by large disparities at high latitudes of both hemispheres, are primarily distinguishable from measurement variability at SH midlatitude and high-latitude surface stations. HIPPO aircraft observations also succeed in differentiating Southern Ocean flux differences from measurement variability. While in situ observations show an ability to discern between ocean flux estimates, remote sensing techniques fail for several reasons. The sparse TCCON network contains no stations below 45°S, the region where flux uncertainty is greatest and where surface observations show the greatest ability to differentiate between ocean flux estimates. As was the case with land flux, ocean flux estimates produce mixing ratio

differences that are smaller than the variability inherent in AIRS data. GOSAT observations, which have smaller single-observation errors than AIRS and greater sensitivity near surface, do not observe ocean regions south of 40°S. The difference between the ocean flux models included in the CMS FPP highlights the larger uncertainty in how high-latitude ocean carbon storage has changed in recent decades and how it may continue to change in response to future climate change. The inability of current remote sensing observations to detect the large differences in ocean flux presented here highlights the need for continued in situ observations and the development of remote sensing techniques which have the potential to increase data yield over high-latitude oceans.

Small mixing ratio differences resulting from flux differences make it difficult to assess the performance of the individual flux estimates. While CASA-GFED compares more favorably than Ames CASA to the total land flux estimate from the Global Carbon Budget [Le Quéré *et al.*, 2013], it still diagnoses the land biosphere as a net source of CO<sub>2</sub> to the atmosphere rather than a sink. NOBM overestimates the magnitude of the ocean sink relative to the Global Carbon Budget and ECCO2-Darwin, but because both land flux estimates underestimate the magnitude of the global sink, the CG-NO simulation produces the most realistic atmospheric growth rate when compared with NOAA surface observations. The main land flux difference that manifests in atmospheric CO<sub>2</sub> mixing ratios is the difference in seasonal cycle between CASA-GFED and Ames CASA. While the earlier spring drawdown in Ames CASA results in slightly better comparisons with surface observations during the spring, earlier release of CO<sub>2</sub> from the land to the atmosphere in autumn degrades comparisons during winter. Small-scale differences between CASA-GFED and Ames CASA, including the presence of fire emissions (CASA-GFED) or redistribution of crop CO<sub>2</sub> (Ames CASA), are difficult to evaluate because these differences are not readily evident in the available observational CO<sub>2</sub> records. While ECCO2-Darwin produces a global ocean sink magnitude nearly equal to the GCB estimate, comparisons with surface and aircraft data show that it is less realistic than NOBM over the Southern Ocean where it diagnoses a larger sink with stronger seasonal variations.

While all flux combinations perform reasonably well in reproducing observed seasonal cycles and spatial gradients in CO<sub>2</sub>, the troubling implication of this agreement between simulations is that even differences between flux models on the order of Pg C are difficult to disentangle using current atmospheric CO<sub>2</sub> observations. Ocean flux differences are particularly difficult to discern because they are smaller and tend to occur in high-latitude regions that are poorly observed by current remote sensing platforms. The use of aircraft campaigns, such as HIPPO, provides a valuable complement to existing long-term carbon monitoring strategies and may become more important in the future as natural carbon reservoirs in remote high-latitude locations respond to a changing climate.

Observations from OCO-2 are expected to greatly improve data yield over much of the globe, which would allow for a greater ability to use space-based observations for flux discrimination. However, OCO-2 will also be limited in its ability to observe high-latitude locations meaning that it will not be able to directly observe many locations associated with large flux uncertainty throughout the year. The Active Sensing of CO<sub>2</sub> emissions over Nights, Days, and Seasons (ASCENDS) mission, which focuses on active rather than passive observations of CO<sub>2</sub>, would improve the ability of GOSAT and OCO-2 to observe high-latitude locations but is not likely to launch until the 2020s. These observational challenges underscore the need for a variety of types of CO<sub>2</sub> observations to help fill the gaps left by satellite observations and to provide additional, complementary information to maximize the impact of satellite observations in regions where they are available.

The small differences in atmospheric CO<sub>2</sub> mixing ratio due to surface fluxes also underscore the importance of quality meteorological analyses and models. The ability to reasonably simulate small gradients in atmospheric CO<sub>2</sub> mixing ratio and to successfully track the transport of air parcels from the surface to the locations and altitudes at which measurements occur is key to the success of inversion studies which seek to reduce uncertainty in natural carbon sinks using atmospheric CO<sub>2</sub> observations. Efforts begun as part of NASA's CMS FPP have succeeded in creating a closer connection between meteorological and CO<sub>2</sub> variations by incorporating CO<sub>2</sub> simulation capabilities into an advanced numerical weather prediction (NWP) model. Similar advances at the European Centre for Medium-Range Weather Forecasts, including the implementation of operational CO<sub>2</sub> forecasts, have been documented by Agustí-Panareda *et al.* [2014]. Closer coordination between the CO<sub>2</sub> and NWP modeling communities will help to minimize and improve quantification of transport model error, ultimately aiding the challenge of improved flux estimation using atmospheric CO<sub>2</sub> observations.

## Acknowledgments

This work was supported by funding from NASA's CMS FPP and Phase One activities. MERRA data used during the CMS FPP have been provided by the GMAO at NASA Goddard Space Flight Center through the NASA GES DISC online archive. Surface observations were provided by NOAA ESRL. TCCON data were obtained from the TCCON Data Archive, operated by the California Institute of Technology from the website at <http://tcon.ipac.caltech.edu/>. HIPPO data were retrieved from the HIPPO data archive at <http://hippo.ornl.gov/dataaccess>. AIRS is managed by NASA's Jet Propulsion Laboratory with data provided through the NASA GES DISC. ACOS data were provided by Chris O'Dell from Colorado State University. Selected monthly average CO<sub>2</sub> fields from the GEOS-5 simulations described above are available at [ftp://gmaoftp.gsfc.nasa.gov/pub/data/lotf/CMS\\_monthly\\_average/](ftp://gmaoftp.gsfc.nasa.gov/pub/data/lotf/CMS_monthly_average/). Additional GEOS-5 CO<sub>2</sub> fields will be provided upon request.

## References

- Adcroft, A., J. Campin, C. Hill, and J. Marshall (2004), Implementation of an atmosphere-ocean general circulation model on the expanded spherical cube, *Mon. Weather Rev.*, **132**, 2845–2863.
- Agusti-Panareda, A., et al. (2014), Forecasting global atmospheric CO<sub>2</sub>, *Atmos. Chem. Phys. Discuss.*, **14**, 13,909–13,962, doi:10.5194/acpd-14-13909-2014.
- Bai, W., X. Zhang, and P. Zhang (2010), Temporal and spatial distribution of tropospheric CO<sub>2</sub> over China based on satellite observations, *Chin. Sci. Bull.*, **55**(31), 3612–3618.
- Bala, G., N. Devaraju, R. K. Chaturvedi, K. Caldeira, and R. Nemani (2013), Nitrogen deposition: How important is it for global terrestrial carbon uptake?, *Biogeosciences*, **10**, 7147–7160, doi:10.5194/bg-10-7147-2013.
- Ballantyne, A. P., C. B. Alden, J. B. Miller, P. P. Tans, and J. W. C. White (2012), Increase in observed net carbon dioxide uptake by land and oceans during the last 50 years, *Nature*, **488**, 70–72.
- Basu, S., et al. (2013), Global CO<sub>2</sub> fluxes estimated from GOSAT retrievals of total column CO<sub>2</sub>, *Atmos. Chem. Phys. Discuss.*, **13**, 4535–4600, doi:10.5194/acpd-13-4535-2013.
- Basu, S., M. Krol, A. Butz, C. Clerbaux, Y. Sawa, T. Machida, H. Matsueda, C. Frankenberg, O. P. Hasekamp, and I. Aben (2014), The seasonal variation of the CO<sub>2</sub> flux over Tropical Asia estimated from GOSAT, CONTRAIL, and IASI, *Geophys. Res. Lett.*, **41**, 1809–1815, doi:10.1002/2013GL059105.
- Bellassen, V., N. Viovy, S. Luyssaert, G. Le Maire, M.-J. Schelhaas, and P. Ciais (2011), Reconstruction and attribution of the carbon sink of European forests between 1950 and 2000, *Global Change Biol.*, **17**, 3274–3292, doi:10.1111/j.1365-2486.2011.02476.x.
- Boden, T. A., G. Marland, and R. J. Andres (2011), *Global, Regional, and National Fossil-Fuel CO<sub>2</sub> Emissions*, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., U.S. Dep. of Energy, Oak Ridge, Tenn., doi:10.3334/CDIAC/00001\_V2011.
- Bousquet, P., P. Ciais, P. Peylin, M. Ramonet, and P. Monfray (1999), Inverse modeling of annual atmospheric CO<sub>2</sub> sources and sinks: 1 Method and control inversion, *J. Geophys. Res.*, **104**(D21), 26,161–26,178, doi:10.1029/1999JD900342.
- Chahine, M. T., L. Chen, P. Dimotakis, X. Jiang, Q. Li, E. T. Olsen, T. Pagano, J. Randerson, and Y. L. Yung (2008), Satellite remote sounding of mid-tropospheric CO<sub>2</sub>, *Geophys. Res. Lett.*, **35**, L17807, doi:10.1029/2008GL035022.
- Chahine, M., C. Barnet, E. T. Olsen, L. Chen, and E. Maddy (2005), On the determination of atmospheric minor gases by the method of vanishing partial derivatives with application to CO<sub>2</sub>, *Geophys. Res. Lett.*, **32**, L22803, doi:10.1029/2005GL024165.
- Chatterjee, A., R. J. Engelen, S. R. Kawa, C. Sweeney, and A. M. Michalak (2013), Background error covariance estimation for atmospheric CO<sub>2</sub> data assimilation, *J. Geophys. Res. Atmos.*, **118**, 10,140–10,154, doi:10.1002/jgrd.50654.
- Chevallier, F., R. J. Engelen, C. Carouge, T. J. Conway, P. Peylin, C. Pickett-Heaps, M. Ramonet, P. J. Rayner, and I. Xueref-Remy (2009), AIRS-based versus flask-based estimation of carbon surface fluxes, *J. Geophys. Res.*, **114**, D20303, doi:10.1029/2009JD012311.
- Chevallier, F., et al. (2011), Global CO<sub>2</sub> fluxes inferred from surface air-sample measurements and from TCCON retrievals of the CO<sub>2</sub> total column, *Geophys. Res. Lett.*, **38**, L24810, doi:10.1029/2011GL049899.
- Chevallier, F., P. I. Palmer, L. Feng, H. Boesch, C. W. O'Dell, and P. Bousquet (2014), Toward robust and consistent regional CO<sub>2</sub> flux estimates from in situ and spaceborne measurements of atmospheric CO<sub>2</sub>, *Geophys. Res. Lett.*, **41**, 1065–1070, doi:10.1002/2013GL058772.
- Ciais, P., P. Bousquet, A. Freibauer, and T. Naegler (2007), Horizontal displacement of carbon associated with agriculture and its impacts on atmospheric CO<sub>2</sub>, *Global Biogeochem. Cycles*, **21**, GB2014, doi:10.1029/2006GB002741.
- Conway, J., P. P. Tans, L. S. Waterman, K. W. Thoning, D. R. Kitzis, K. A. Masarie, and N. Zhang (1994), Evidence of interannual variability of the carbon cycle from the NOAA/CMDL global air sampling network, *J. Geophys. Res.*, **99**(D11), 22,831–22,855, doi:10.1029/94JD01951.
- Crisp, D., et al. (2012), The ACOS CO<sub>2</sub> retrieval algorithm—Part II: Global XCO<sub>2</sub> data characterization, *Atmos. Meas. Tech.*, **5**, 687–707, doi:10.5194/amt-5-687-2012.
- Doney, S. C., et al. (2004), Evaluating global ocean carbon models: The importance of realistic physics, *Global Biogeochem. Cycles*, **18**, GB3017, doi:10.1029/2003GB002150.
- Dutkiewicz, S., M. Follows, and J. Bragg (2009), Modeling the coupling of ocean ecology and biogeochemistry, *Global Biogeochem. Cycles*, **23**, GB4017, doi:10.1029/2008GB003405.
- Engelen, R. J., S. Serran, and F. Chevallier (2009), Four-dimensional data assimilation of atmospheric CO<sub>2</sub> using AIRS observations, *J. Geophys. Res.*, **114**, D03303, doi:10.1029/2008JD010739.
- Esser, G., J. Kattge, and A. Sakalli (2011), Feedback of carbon and nitrogen cycles enhances carbon sequestration in the terrestrial biosphere, *Global Change Biol.*, **17**, 819–842, doi:10.1111/j.1365-2486.2010.02261.x.
- Feng, L., P. I. Palmer, Y. Yang, R. M. Yantosca, S. R. Kawa, J.-D. Paris, H. Matsueda, and T. Machida (2011), Evaluating a 3-D transport model of atmospheric CO<sub>2</sub> using ground-based, aircraft, and space-borne data, *Atmos. Chem. Phys.*, **11**, 2789–2803, doi:10.5194/acp-11-2789-2011.
- Fleischer, K., et al. (2013), The contribution of nitrogen deposition to the photosynthetic capacity of forests, *Global Biogeochem. Cycles*, **27**, 187–199, doi:10.1002/gbc.20026.
- Follows, M., and S. Dutkiewicz (2011), Modeling diverse communities of marine microbes, *Annu. Rev. Mar. Sci.*, **3**, 427–451.
- Follows, M., T. Ito, and S. Dutkiewicz (2006), On the solution of the carbonate chemistry system in ocean biogeochemistry models, *Ocean Modell.*, **12**, 290–301.
- Follows, M., S. Dutkiewicz, S. Grant, and S. Chisholm (2007), Emergent biogeography of microbial communities in a model ocean, *Science*, **315**, 1843–1846.
- Gerber, S., L. O. Hedin, S. G. Keel, S. W. Pacala, and E. Shevliakova (2013), Land use change and nitrogen feedbacks constrain the trajectory of the land carbon sink, *Geophys. Res. Lett.*, **40**, 5218–5222, doi:10.1002/grl.50957.
- Gourdji, S. M., A. I. Hirsch, K. L. Mueller, V. Yadav, A. E. Andrews, and A. M. Michalak (2010), Regional-scale geostatistical inverse modeling of North American CO<sub>2</sub> fluxes: A synthetic data study, *Atmos. Chem. Phys.*, **10**, 6151–6167, doi:10.5194/acp-10-6151-2010.
- Gregg, W. W., and N. W. Casey (2007), Modeling coccolithophores in the global oceans, *Deep Sea Res., Part II*, **54**, 447–477.
- Gregg, W. W., P. Ginoux, P. S. Schopf, and N. W. Casey (2003), Phytoplankton and Iron: Validation of a global three-dimensional ocean biogeochemical model, *Deep Sea Res., Part II*, **50**, 3143–3169.
- Gregg, W. W., N. W. Casey, and C. S. Rousseaux (2013), *Global Surface Ocean Estimates in a Model Forced by MERRA*, NASA Global Model. and Assimilation Ser., vol. 31, edited by M. Suarez, *Tech. Rep.*, 104606 V31, NASA.
- Guerlet, S., S. Basu, A. Butz, M. Krol, P. Hahne, S. Houweling, O. P. Hasekamp, and I. Aben (2013), Reduced carbon uptake during the 2010 Northern Hemisphere summer from GOSAT, *Geophys. Res. Lett.*, **40**, 2378–2383, doi:10.1002/grl.50402.
- Gurney, K. R., et al. (2002), Towards robust regional estimates of CO<sub>2</sub> sources and sinks using atmospheric transport models, *Nature*, **415**, 626–630.
- Heimbach, P., D. Menemenlis, M. Losch, J. Campin, and C. Hill (2010), On the formulation of sea-ice models. Part 2: Lessons from multi-year adjoint sea ice export sensitivities through the Canadian Arctic Archipelago, *Ocean Modell.*, **33**, 145–158.



- Hollinger, D. Y., et al. (2004), Spatial and temporal variability in forest-atmosphere CO<sub>2</sub> exchange, *Global Change Biol.*, 10, 1689–1706, doi:10.1111/j.1365-2486.2004.00847.x.
- Huntzinger, D. N., S. M. Gourdji, K. L. Mueller, and A. M. Michalak (2011), The utility of continuous atmospheric measurements for identifying biospheric CO<sub>2</sub> flux variability, *J. Geophys. Res.*, 116, D06110, doi:10.1029/2010JD015048.
- Huntzinger, D. N., et al. (2012), North American Carbon Program (NACP) regional interim synthesis: Terrestrial biospheric model intercomparison, *Ecol. Modell.*, 232, 144–157.
- Inoue, M., et al. (2013), Validation of XCO<sub>2</sub> derived from SWIR spectra of GOSAT TANSO-FTS with aircraft measurement data, *Atmos. Chem. Phys. Discuss.*, 13, 3203–3246, doi:10.5194/acpd-13-3203-2013.
- Jiang, X., M. T. Chahine, E. T. Olsen, L. L. Chen, and Y. L. Yung (2010), Interannual variability of midtropospheric CO<sub>2</sub> from Atmospheric Infrared Sounder, *Geophys. Res. Lett.*, 37, L13801, doi:10.1029/2010GL042823.
- Keppel-Aleks, G., et al. (2012), The imprint of surface fluxes and transport on variations in total column carbon dioxide, *Biogeosciences*, 9, 875–891, doi:10.5194/bg-9-875-2012.
- Keppel-Aleks, G., et al. (2013), Atmospheric carbon dioxide variability in the Community Earth System Model: Evaluation and transient dynamic during the 20th and 21st centuries, *J. Clim.*, 26, 4447–4475, doi:10.1175/JCLI-D-12-00589.1.
- Kipling, Z., P. Stier, J. P. Schwarz, A. E. Perring, J. R. Spackman, G. W. Mann, C. E. Johnson, and P. J. Telford (2013), Constraints on aerosol processes in climate models from vertically-resolved aircraft observations of black carbon, *Atmos. Chem. Phys.*, 13, 5969–5986, doi:10.5194/acp-13-5969-2013.
- Kort, A. E., C. Frankenberg, C. E. Miller, and T. Oda (2012), Space-based observations of megacity carbon dioxide, *Geophys. Res. Lett.*, 39, L17806, doi:10.1029/2012GL052738.
- Kuze, A., S. Hiroshi, N. Masakatsu, and H. Takashi (2009), Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, *Appl. Opt.*, 48, 6716–6733, doi:10.1364/AO.48.006716.
- Lafleur, P. M., N. T. Roulet, J. L. Bubier, S. Frolking, and T. R. Moore (2003), Interannual variability in the peatland-atmosphere carbon dioxide exchange at an ombrotrophic bog, *Global Biogeochem. Cycles*, 17(2), 1036, doi:10.1029/2002GB001983.
- Law, R. M., et al. (2008), TransCom model simulations of hourly atmospheric CO<sub>2</sub>: Experimental overview and diurnal cycle results for 2002, *Global Biogeochem. Cycles*, 22, GB3009, doi:10.1029/2007GB003050.
- Le Quéré, C., et al. (2013), The global carbon budget 1959–2011, *Earth Syst. Sci. Data*, 5, 165–185, doi:10.5194/essd-5-165-2013.
- Li, K. F., B. Tian, D. E. Waliser, and Y. L. Yung (2010), Tropical mid-tropospheric CO<sub>2</sub> variability driven by the Madden-Julian oscillation, *Proc. Natl. Acad. Sci. U.S.A.*, 107(45), 19,171–19,175, doi:10.1073/pnas.1008222107.
- Lin, S.-J. (2004), A “vertically Lagrangian” finite-volume dynamical core for global models, *Mon. Weather Rev.*, 132(10), 2293–2307.
- Liu, J., I. Fung, E. Kalnay, J.-S. Kang, E. T. Olsen, and L. Chen (2012), Simultaneous assimilation of AIRS XCO<sub>2</sub> and meteorological observations in a carbon climate model with an ensemble Kalman filter, *J. Geophys. Res.*, 117, D05309, doi:10.1029/2011JD016642.
- Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith (2000), A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests, *Mon. Weather Rev.*, 128, 3187–3199.
- Los, S. O. (2013), Analysis of trends in fused AVHRR and MODIS NDVI data for 1982–2006: Indication for a CO<sub>2</sub> fertilization effect in global vegetation, *Global Biogeochem. Cycles*, 27, 318–330, doi:10.1002/gbc.20027.
- Los, S. O., G. J. Collatz, P. J. Sellers, C. M. Malmstrom, N. H. Pillar, R. S. DeFries, L. Bounoua, M. T. Parris, C. J. Tucker, and D. A. Dazlich (2000), A global 9-yr biophysical land surface dataset from NOAA AVHRR data, *J. Hydrometeorol.*, 1, 183–199.
- Losch, M., D. Menemenlis, P. Heimbach, J. Campin, and C. Hill (2010), On the formulation of sea-ice models. Part 1: Effects of different solver implementations and parameterizations, *Ocean Modell.*, 33, 129–144.
- Maksyutov, S., et al. (2013), Regional CO<sub>2</sub> flux estimates for 2009–2010 based on GOSAT and ground-based CO<sub>2</sub> observations, *Atmos. Chem. Phys.*, 13, 9351–9373, doi:10.5194/acp-13-9351-2013.
- Mann, G. W., et al. (2012), Intercomparison of modal and sectional aerosol microphysics representations within the same 3-D global chemical transport model, *Atmos. Chem. Phys.*, 12, 4449–4476, doi:10.5194/acp-12-4449-2012.
- Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey (1997a), A finite-volume, incompressible Navier-Stokes model for studies of the ocean on parallel computers, *J. Geophys. Res.*, 102(C3), 5753–5766, doi:10.1029/96JC02775.
- Marshall, J., C. Hill, L. Perelman, and A. Adcroft (1997b), Hydrostatic, quasi-hydrostatic and non-hydrostatic ocean modeling, *J. Geophys. Res.*, 102(C3), 5733–5752, doi:10.1029/96JC02776.
- Menemenlis, D., et al. (2005a), NASA supercomputer improves prospects for ocean climate research, *Eos Trans. AGU*, 86(89), 95–96.
- Menemenlis, D., I. Fukumori, and T. Lee (2005b), Using Green’s functions to calibrate an ocean general circulation model, *Mon. Weather Rev.*, 133, 1224–1240.
- Menemenlis, D., J. Campin, P. Heimbach, C. Hill, T. Lee, A. Nguyen, M. Schodlock, and H. Zhang (2008), ECCO2: High resolution global ocean and sea ice data synthesis, *Mercator Ocean Quart. Newslett.*, 31, 13–21.
- Messerschmidt, J., et al. (2012), Automated ground-based remote sensing measurements of greenhouse gases at the Białystok site in comparison with collocated in situ measurements and model data, *Atmos. Chem. Phys.*, 12(15), 6741–6755, doi:10.5194/acp-12-6741-2012.
- Michalak, A. M., L. Bruhwiler, and P. P. Tans (2004), A geostatistical approach to surface flux estimation of atmospheric trace gases, *J. Geophys. Res.*, 109, D14109, doi:10.1029/2003JD004422.
- Moorthi, S., and M. J. Suarez (1992), Relaxed Arakawa-Schubert: A parameterization of moist convection for general circulation models, *Mon. Weather Rev.*, 120, 978–1002.
- Mu, M., et al. (2011), Daily and 3-hourly variability in global fire emissions and consequences for atmospheric model predictions of carbon monoxide, *J. Geophys. Res.*, 116, D24303, doi:10.1029/2011JD016245.
- Najjar, R. G., et al. (2007), Impact of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean: Results from Phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2), *Global Biogeochem. Cycles*, 21, GB3007, doi:10.1029/2006GB002857.
- Novelli, P. C., L. P. Steele, and P. P. Tans (1992), Mixing ratios of carbon monoxide in the troposphere, *J. Geophys. Res.*, 97(D18), 20,731–20,750, doi:10.1029/92JD02010.
- O’Dell, C. W., et al. (2012), The ACOS CO<sub>2</sub> retrieval algorithm—Part 1: Description and validation against synthetic observations, *Atmos. Meas. Tech.*, 5, 99–121, doi:10.5194/amt-5-99-2012.
- Olsen, S. C., and J. T. Randerson (2004), Differences between surface and column atmospheric CO<sub>2</sub> and implications for carbon cycle research, *J. Geophys. Res.*, 109, D02301, doi:10.1029/2003JD003968.
- Orr, J. C., et al. (2001), Estimates of anthropogenic carbon uptake from four three-dimensional global ocean models, *Global Biogeochem. Cycles*, 15(1), 43–60, doi:10.1029/2000GB001273.

- Ott, L., B. Duncan, S. Pawson, P. Colarco, M. Chin, C. Randles, T. Diehl, and J. E. Nielsen (2010), Influence of the 2006 Indonesian biomass burning aerosols on tropical dynamics studied with the GEOS-5 AGCM, *J. Geophys. Res.*, **115**, D14121, doi:10.1029/2009JD013181.
- Ott, L., S. Pawson, and J. Bacmeister (2011), An analysis of the impact of convective parameter sensitivity on simulated global atmospheric CO distributions, *J. Geophys. Res.*, **116**, D21310, doi: 10.1029/2011JD016077.
- Piao, S., et al. (2013), Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO<sub>2</sub> trends, *Global Change Biol.*, **19**, 2117–2132, doi:10.1111/gcb.12187.
- Pinzon, J. E., and C. J. Tucker (2014), A non-stationary 1981–2012 AVHRR NDVI3g time series, *Remote Sens.*, **6**, 6929–6960.
- Popova, E. E., A. Yool, A. C. Coward, F. Dupont, C. Deal, S. Elliott, E. Hunke, M. Jin, M. Steele, and J. Zhang (2012), What controls primary production in the Arctic Ocean? Results from an intercomparison of five general circulation models with biogeochemistry, *J. Geophys. Res.*, **117**, C00D12, doi:10.1029/2011JC007112.
- Potter, C. S. (1997), An ecosystem simulation model for methane production and emission from wetlands, *Global Biogeochem. Cycles*, **11**(4), 495–506, doi:10.1029/97GB02302.
- Potter, C. S., J. T. Randerson, C. B. Field, P. A. Matson, P. M. Vitousek, H. A. Mooney, and S. A. Klooster (1993), Terrestrial ecosystem production—A process model based on global satellite and surface data, *Global Biogeochem. Cycles*, **7**(4), 811–841, doi:10.1029/93GB02725.
- Potter, C., S. Klooster, A. Huete, and V. Genovese (2007), Terrestrial carbon sinks for the United States predicted from MODIS satellite data and ecosystem modeling, *Earth Interact.*, **11**, 1–21, doi:10.1175/EI228.1.
- Potter, C., S. Klooster, V. Genovese, C. Hiatt, S. Boriah, V. Kumar, V. Mithal, and A. Garg (2012), Terrestrial ecosystem carbon fluxes predicted from MODIS satellite data and large-scale disturbance modeling, *Int. J. Geosci.*, doi:10.4236/ijg.2012.
- Raczka, B. M., et al. (2013), Evaluation of continental carbon cycle simulations with North American flux tower observations, *Ecol. Monogr.*, **83**(4), 531–556.
- Randerson, J. T., M. V. Thompson, and C. M. Malmstrom (1996), Substrate limitations for heterotrophs: Implications for models that estimate the seasonal cycle of atmospheric CO<sub>2</sub>, *Global Biogeochem. Cycles*, **10**(4), 585–602, doi:10.1029/96GB01981.
- Randerson, J. T., et al. (2009), Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, *Global Change Biol.*, **15**, 2462–2484, doi:10.1111/j.1365-2486.2009.01912.
- Reinecker, M. M., et al. (2008), The GEOS-5 Data Assimilation System—Documentation of Versions 5.0.1, 5.1.0, and 5.2.0, Tech. Rep. 104606 V27, NASA.
- Rienecker, M. M., et al. (2011), MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Clim.*, **24**, 3624–3648.
- Ruzmaikin, A., H. H. Aumann, and T. S. Pagano (2012), Patterns of CO<sub>2</sub> variability from global satellite data, *J. Clim.*, doi:10.1175/JCLI-D-11-00223.1.
- Schaefer, K., et al. (2012), A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis, *J. Geophys. Res.*, **117**, G03010, doi:10.1029/2012JG001960.
- Schuh, A. E., A. S. Denning, K. D. Corbin, I. T. Baker, M. Uliasz, N. Parazoo, A. E. Andrews, and D. E. J. Worthy (2010), A regional high-resolution carbon flux inversion of North America for 2004, *Biogeosciences*, **7**(5), 1625–1644, doi:10.5194/bg-7-1625-2010.
- Schwalm, C. R., et al. (2010), A model-data intercomparison of CO<sub>2</sub> exchange across North America: Results from the North American Carbon Program site synthesis, *J. Geophys. Res.*, **115**, G00H05, doi:10.1029/2009JG001229.
- Sentman, L. T., E. Shevliakova, R. J. Stouffer, and S. Malyshev (2011), Time scales of terrestrial carbon response related to land-use application: Implications for initializing an Earth system model, *Earth Interact.*, **15**, 1–16, doi:10.1175/2011EI401.1.
- Takagi, H., et al. (2011), On the benefit of GOSAT observations to the estimation of regional CO<sub>2</sub> fluxes, *SOLA*, **7**, 161–164, doi:10.2151/sola.2011-041.
- Takahashi, T., et al. (2002), Global sea-air CO<sub>2</sub> flux based on climatological surface ocean pCO<sub>2</sub>, and seasonal biological and temperature effects, *Deep Sea Res., Part II*, **49**, 1601–1622.
- Takahashi, T., et al. (2009), Climatological mean and decadal changes in surface ocean pCO<sub>2</sub>, and net sea-air CO<sub>2</sub> flux over the global oceans, *Deep Sea Res., Part II*, **56**, 554–577.
- Tans, P. P., I. Y. Fung, and T. Takahashi (1990), Observational constraints on the global atmospheric CO<sub>2</sub> budget, *Science*, **247**, 1431–1438.
- Tucker, C. J., J. E. Pinzon, M. E. Brown, D. A. Slayback, E. W. Pak, R. Mahoney, E. F. Vermote, and N. El Saleous (2005), An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, *Int. J. Remote Sens.*, **26**(20), 4485–4498.
- van der Werf, G. R., J. T. Randerson, G. J. Collatz, L. Giglio, P. S. Kasibhatla, A. F. Arellano, S. C. Olsen, and E. S. Kasischke (2004), Continental-scale partitioning of fire emissions during the 1997 to 2001 El Niño/La Niña period, *Science*, **303**, 73–76.
- van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano (2006), Interannual variability of global biomass burning emissions from 1997 to 2004, *Atmos. Chem. Phys.*, **6**, 3423–3441.
- van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, M. Mu, P. S. Kasibhatla, D. C. Morton, R. S. DeFries, Y. Jin, and T. T. van Leeuwen (2010), Global fire emissions and the contribution of deforestation, agriculture, and peat fires (1997–2009), *Atmos. Chem. Phys.*, **10**, 11,707–11,735.
- Vanuytrecht, E., D. Raes, and P. Willems (2011), Considering sink strength to model crop production under elevated atmospheric CO<sub>2</sub>, *Agric. Forest Meteorol.*, **151**, 1753–1762, doi:10.1016/j.agrformet.2011.07.011.
- Wang, H., D. J. Jacob, M. Kopacz, D. B. A. Jones, P. Suntharalingam, J. A. Fisher, R. Nassar, and S. Pawson (2009), Error correlation between CO<sub>2</sub> and CO as a constraint for CO<sub>2</sub> flux inversions using satellite data, *Atmos. Chem. Phys.*, **9**, 7313–7323, doi:10.5194/acp-9-7313-2009.
- Wanninkhof, R. (1992), Relationship between wind speed and gas exchange over the ocean, *J. Geophys. Res.*, **97**(C5), 7373–7382, doi:10.1029/92JC00188.
- Winderlich, J., C. Gerbig, O. Kolle, and M. Heimann (2014), Inferences from CO<sub>2</sub> and CH<sub>4</sub> concentration profiles at the Zotino Tall Tower Observatory (ZOTTO) on regional summertime ecosystem fluxes, *Biogeosciences*, **11**, 2055–2068, doi:10.5194/bg-11-2055-2014.
- Wofsy, S. C., et al. (2011), HIPPO Pole-to-Pole Observations (HIPPO): Fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, *Proc. R. Soc. A*, **369**, 2073–2086, doi:10.1098/rsta.2010.0313.
- Wunch, D., et al. (2010), Calibration of the total carbon column observing network using aircraft profile data, *Atmos. Measure. Tech.*, **3**(5), 1351–1362, doi:10.5194/amt-3-1351-2010.
- Wunch, D., G. C. Toon, J.-F. L. Blavier, R. A. Washenfelder, J. Notholt, B. J. Connor, D. W. T. Griffith, V. Sherlock, and P. O. Wennberg (2011a), The total carbon column observing network, *Philos. Trans. R. Soc. London, Ser. A*, **369**(1943), 2087–2112, doi:10.1098/rsta.2010.0240.
- Wunch, D., et al. (2011b), A method for evaluating bias in global measurements of CO<sub>2</sub> total columns from space, *Atmos. Chem. Phys.*, **11**(23), 12,317–12,337, doi:10.5194/acp-11-12317-2011.

Wunsch, C., and P. Heimbach (2007), Practical global ocean state estimation, *Phys. D*, *230*, 197–208.

Yokota, T., Y. Yoshida, N. Eguchi, Y. Ota, T. Tanaka, H. Watanabe, and S. Maksyutov (2009), Global concentrations of CO<sub>2</sub> and CH<sub>4</sub> retrieved from GOSAT: First preliminary results, *SOLA*, *5*, 160–163.

Zhu, Y., and R. Gelaro (2008), Observation sensitivity calculations using the adjoint of the Gridpoint Statistical Interpolation (GSI) analysis system, *Mon. Weather Rev.*, *136*, 335–351.